

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 7 :</b> A01K 67/00, A01N 61/00, 63/00, A61K 38/00, 48/00, C12N 5/00, 5/02, C12P 21/00, C12Q 1/00	<b>A1</b>	<b>(11) International Publication Number:</b> WO 00/32039 <b>(43) International Publication Date:</b> 8 June 2000 (08.06.00)
<b>(21) International Application Number:</b> PCT/US99/28764 <b>(22) International Filing Date:</b> 3 December 1999 (03.12.99)  <b>(30) Priority Data:</b> 60/110,901 4 December 1998 (04.12.98) US  <b>(71) Applicant (for all designated States except US):</b> DUKE UNIVERSITY [US/US]; 230 North Building, Research Drive, Durham, NC 27708 (US).  <b>(72) Inventor; and</b> <b>(75) Inventor/Applicant (for US only):</b> LIN, Haifan [US/US]; 2620 McDowell Road, Durham, NC 27705 (US).  <b>(74) Agent:</b> TAYLOR, Arles, A., Jr.; Jenkins & Wilson, P.A., Suite 1400, University Tower, 3100 Tower Boulevard, Durham, NC 27707 (US).		<b>(81) Designated States:</b> AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> PURIFIED AND ISOLATED <i>piwi</i> FAMILY GENES AND GENE PRODUCTS AND METHODS EMPLOYING SAME  <b>(57) Abstract</b>  A family of isolated and purified proteins and nucleic acids are disclosed. Particularly, <i>piwi</i> family proteins and cDNAs encoding the same are disclosed. Recombinant host cells, recombinant nucleic acids, recombinant proteins and transgenic animals are also disclosed, along with methods of producing each. Isolated and purified antibodies to <i>piwi</i> family homologs, and methods of producing the same, are also disclosed. <i>piwi</i> family gene products are characterized as having activity in the growth, proliferation and self-renewing division of stem cells, and proliferation of primordial germ cells. Thus, therapeutic, screening, culturing and transgenic methods involving these activities are also disclosed.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Description

PURIFIED AND ISOLATED *piwi* FAMILY GENES AND GENE PRODUCTS  
AND METHODS EMPLOYING SAME

Priority Application Information

5           This application claims the benefit of United States Provisional Patent  
Application 60/110,091, filed December 4, 1998, the disclosure of which is  
incorporated herein by reference in its entirety.

Grant Statement

          This work was supported by NIH grant HD 33760. The U.S.  
10       Government has certain rights in the invention.

Technical Field

          The present invention relates generally to isolated and purified proteins  
and nucleic acids which modulate stem cell renewal, growth and division and  
which modulate primordial germ cell proliferation. More particularly, the  
15       present invention relates to isolated and purified *piwi* family proteins and  
isolated and purified polynucleic acids encoding the same.

          The publications and other materials used herein to illuminate the  
background of the invention, and in particular cases, to provide additional  
details respecting the practice, are incorporated herein by reference, and for  
20       convenience, are referenced by author and year of publication in the following  
text, and respectively grouped by author in the appended list of references.

Table of Abbreviations

APC	apical polar cell(s)
ATCC	American Type Culture Collection
25   bFGF	basic fibroblast growth factor
BRL	buffalo rat liver
BSA	bovine serum albumin
Cb	cystoblast

-2-

	CDR(s)	complementarity determining region(s)
	COOH	carboxy terminus
	CyO	a strain of <i>Drosophila</i>
	DAPI	a DNA specific dye
5	DMEM	Dulbecco's modified Eagle's medium
	DC	differentiating 16-cell cysts
	DTC	distal tip cell
	ECL	enhanced chemiluminescent
	EDTA	ethylenediamine tetraacetate
10	ES	epithelial sheath
	EST	expressed sequence tags
	FBS	fetal bovine serum
	FC	follicle cell
	<i>FRT40A</i>	a strain of <i>Drosophila</i>
15	Fu	fusome
	GC-MS	gas chromatography-mass spectroscopy
	Ge	germarium(ia)
	GSC(s)	germline stem cell(s)
	HAT	cell culture media comprising hypoxanthine, aminopterin, and thymidine
20	<i>hiwi/Hiwi</i>	a human homolog of the <i>piwi/PIWI</i> gene and gene product
	HPLC	high pressure liquid chromatography
	hsGal4	a heat shock inducible transgene
25	IS	inner sheath
	KLH	keyhole limpet hemocyanin
	MC	mitotically active cysts
	<i>miwi/MIWI</i>	a murine homolog of the <i>piwi/PIWI</i> gene and gene product
30	MPZ	mitotic proliferation zone
	myc	human oncogene used herein as molecular tag for <i>piwi</i>



-3-

	NH2	amino terminus
	ORF	open reading frame
	PCR	polymerase chain reaction
	PGC(s)	primordial germ cell(s)
5	<i>piwi</i> /PIWI	a gene and gene product isolated from <i>Drosophila</i> and having stem cell division modulating activity
	RACE	rapid amplification of cDNA ends
	RNAi	RNA interference assay
	SSC(s)	somatic stem cell(s)
10	Sp	spectrosome
	TFC	terminal filament cells
	UTR	untranslated region
	VASA	a germline specific marker
	WT	wild type

15

**Background Art**

Stem cells are a very small number of founder cells that play a central role in tissue development and maintenance. In human bodies, stem cells are responsible for generating and/or maintaining approximately 90% of cells in the adult tissues. Over-proliferation of malignant stem cells is the leading cause of cancer while under-proliferation of stem cells or stem-like progenitor cells leads to tissue dystrophy, anemia, immunodeficiency, and male infertility. The crucial role of stem cells has long been attributed to their ability to self-renew and to generate immense number of specialized cells on demand.

The ability of stem cells to self-renew and to produce a large number of differentiated progeny is critical for the development and maintenance of a wide variety of tissues in organisms ranging from insects to mammals (reviewed in Potten, 1997; Lin, 1997; Lin and Schagat, 1997; Morrison et al., 1997). This self-renewing ability is controlled both by extrinsic signaling and by cell-autonomous mechanisms (reviewed in Morrison et al., 1997; Lin and Schagat, 1997). Cell autonomous mechanisms have been elucidated in a few stem cell models such as neuroblasts and germline stem cells in *Drosophila* (Lin and

Schagat, 1997; Deng and Lin, 1997), whereas the role of extrinsic signaling has been elucidated in several systems. For example, the proliferation and differentiation of mammalian stem cells in the hematopoietic, epidermal, and nervous systems depend on extrinsic signals that act on specific receptors on the stem cell surface (Morrison et al., 1997).

In diverse organisms ranging from invertebrates to mammals, the proliferation of germ cells, some of which possess stem cell properties, has been postulated, and, in some cases, shown to be regulated by neighboring non-mitotic somatic cells (Lin, 1997). Particularly, in *C. elegans*, cell-cell interactions between the somatic distal tip cell (DTC) at the end of each gonadal arm and the underlying mitotic germline nuclei via the *lag-2/glp-1* signaling pathway provides a paradigm for soma-germline interaction (reviewed in Kimble and Simpson, 1997). The *glp-1* pathway is required to maintain a population of mitotically active nuclei in the germline.

However, few molecules and/or mechanisms identified in a particular type of stem cells have been shown to be applicable to other stem cell systems. For example, the *glp-1* equivalent pathway in *Drosophila* does not play a role in regulating GSC division and maintenance (Ruohala et al., 1991; Xu et al., 1992).

The self-renewing asymmetric division of GSCs in the *Drosophila* ovary is known to be controlled both by an intracellular mechanism (Deng and Lin, 1997) and by cell-cell interactions (Lin and Spradling, 1993). The intracellular mechanism involves a cytoplasmic organelle termed the spectrosome that controls the orientation of GSC division (Lin et al., 1994; Deng and Lin, 1997). The cell-cell interaction mechanism involves terminal filament cells, as shown by laser ablation studies (Lin and Spradling, 1993). Recently, *dpp* has been shown as a key signaling molecule required for GSC division and maintenance (Xie and Spradling, 1998). It is possible that the *dpp* signal emanates from somatic cells. Alternatively, *dpp* signal may originate from the germline or even within GSCs, like its mammalian homologs (Zhao et al., 1996).

In mammals, primordial germ cells cultured from the genital ridge have the ability to give rise to pluripotent embryonic stem cells. For example, U.S.

-5-

Patent No. 5,690,926 issued November 25, 1997 to Hogan; U.S. Patent No. 5,670,372 issued September 23, 1997 to Hogan; and U.S. Patent No. 5,537,357 issued September 26, 1995 to Hogan each disclose pluripotent mammalian embryonic stem cells and methods of making the same. The disclosure of these patents is limited to mammalian embryonic stem cells and particularly to the culturing of murine and other mammalian embryonic stem cells using a combination of growth factors consisting of SCF, FGF and LIF.

Current prior art reports on the culture of avian primordial germ cells (PGCs) have concentrated on efforts to maintain a PGC-phenotype and to stimulate proliferation. See e.g., Chang, I. K. et al., *Cell. Biol. Int.* 1997 Aug; 21(8): 495-9; Chang, I. K. et al., *Cell. Biol. Int.* 1995 Feb; 19(2): 143-9; Allioli, N. et al., *Dev. Biol.* 1994 Sep; 165(1): 30-7 and PCT Publication No. WO 99/06533, published February 11, 1999 (Applicant - University of Massachusetts; Inventors - Ponce de Leon et al.).

As illustrated above, numerous attempts have been devoted to identify genes that control the self-renewing ability of stem cells or the proliferation of primordial germ cells. As a result, a number of growth factors and signaling molecules, such as *Steel* factor and its *c-kit* receptor, have been identified to regulate such activity in certain tissues. Despite this progress, there remains a long-felt and continuing need to identify genes that play a role in modulating the growth and self-renewing division of stem cells, particularly GSCs, and that play a role in modulating proliferation of primordial germ cells.

#### Summary of the Invention

The present invention contemplates an isolated and purified family of genes and gene products (the *piwi* family) which plays a role in the growth, proliferation and self-renewing division of stem cells, and proliferation of primordial germ cells. More preferably, a polypeptide of the invention is a recombinant polypeptide. Even more preferably, a polypeptide of the present invention comprises a vertebrate *piwi* family polypeptide. Even more preferably, a polypeptide of the present invention comprises a mammalian *piwi* family polypeptide. Even more preferably, a polypeptide of the present

-6-

invention comprises a human *piwi* family polypeptide. Even more preferably, a polypeptide of the present invention comprises an amino acid sequence from the amino acid residue sequences of any of SEQ ID NOs:2, 4 and 6.

5 The present invention also provides an isolated and purified polynucleotide that encodes a polypeptide that plays a role in the growth, proliferation and self-renewing division of stem cells, and proliferation of primordial germ cells. In a preferred embodiment, a polynucleotide of the present invention comprises a DNA molecule from a vertebrate species. Preferred vertebrates comprise mammals, birds or fish. A preferred mammal  
10 is a human. More preferably, a polynucleotide of the present invention encodes a polypeptide designated PIWI. Even more preferably, a polynucleotide of the present invention encodes a polypeptide comprising an amino acid residue sequence of any of SEQ ID NOs:2, 4 and 6. Most preferably, an isolated and purified polynucleotide of the invention comprises  
15 a nucleotide base sequence of any of SEQ ID NOs:1, 3 and 5.

In another embodiment, the present invention contemplates an antibody immunoreactive with a *piwi* family polypeptide as described above. SEQ ID NOs:3-6 set forth nucleotide and amino acid sequences from representative vertebrates, human and mouse. Also contemplated by the present invention  
20 are antibodies immunoreactive with homologues or biologically equivalent *piwi* family polynucleotides and polypeptides found in other vertebrates. Preferably, an antibody of the invention is a monoclonal antibody. More preferably, the *piwi* family polypeptide comprises human PIWI (HIWI). Even more preferably, the *piwi* polypeptide comprises an amino acid residue sequence of any of SEQ  
25 ID NOs:2, 4 and 6.

In another aspect, the present invention contemplates a method of producing an antibody immunoreactive with a *piwi* family polypeptide as described above, the method comprising the steps of (a) transfecting a recombinant host cell with a polynucleotide that encodes a biologically active  
30 *piwi* family polypeptide; (b) culturing the host cell under conditions sufficient for expression of the polypeptide; (c) recovering the polypeptide; and (d) preparing the antibody to the polypeptide. SEQ ID NOs:3-6 set forth nucleotide and

-7-

amino acid sequences from representative vertebrates, human and mouse. Preferably, the host cell is transfected with a polynucleotide of any of SEQ ID NOs:1, 3 and 5. Even more preferably, the present invention provides an antibody prepared according to the method described above. Also contemplated by the present invention is the use of homologues or biologically equivalent polynucleotides and polypeptides found in other vertebrates to produce antibodies.

Alternatively, the present invention provides a method of detecting a *piwi* family polypeptide as described above, wherein the method comprises immunoreacting the polypeptide with an antibody prepared according to the method described above to form an antibody-polypeptide conjugate, and detecting the conjugate.

In yet another embodiment, the present invention contemplates a method of detecting a messenger RNA transcript that encodes a *piwi* family polypeptide as described above, wherein the method comprises hybridizing the messenger RNA transcript with a polynucleotide sequence that encodes that polypeptide to form a duplex; and detecting the duplex. Alternatively, the present invention provides a method of detecting a DNA molecule that encodes a *piwi* family polypeptide as described above, wherein the method comprises hybridizing DNA molecules with a polynucleotide that encodes a biologically active *piwi* family polypeptide to form a duplex; and detecting the duplex.

In another aspect, the present invention contemplates an assay kit for detecting the presence of a *piwi* family polypeptide in a biological sample, where the kit comprises a first container containing a first antibody capable of immunoreacting with a biologically active *piwi* polypeptide, with the first antibody. Preferably, the first antibody is present in an amount sufficient to perform at least one assay. Also preferably, an assay kit of the invention further comprises a second container containing a second antibody that immunoreacts with the first antibody. More preferably, the antibodies used in an assay kit of the present invention are monoclonal antibodies. Even more preferably, the first antibody is affixed to a solid support. More preferably still,

the first and second antibodies comprise an indicator, and, preferably, the indicator is a radioactive label, a fluorescent label or an enzyme.

In an alternative aspect, the present invention provides an assay kit for detecting the presence, in biological samples, of a *piwi* family polypeptide, the kits comprising a first container that contains a second polynucleotide identical or complementary to a segment of at least 10 contiguous nucleotide bases of a polynucleotide that encodes a biologically active *piwi* family polypeptide.

In another embodiment, the present invention contemplates an assay kit for detecting the presence, in a biological sample, of an antibody immunoreactive with a *piwi* family polypeptide, the kit comprising a first container containing a biologically active *piwi* family polypeptide that immunoreacts with the antibody, with the polypeptide present in an amount sufficient to perform at least one assay.

In still a further embodiment, this invention pertains to transgenic, therapeutic, screening and culturing methods based upon the self-renewing division of stem cells, and proliferation of primordial germ cells function of *piwi* family polynucleotides and polypeptides as described herein. Representative therapeutic and culturing methods include administration of a soluble form of the *piwi* protein and gene therapy approaches using an isolated and purified polynucleotide of the present invention.

Thus, a key aspect of this invention pertains to the discovery of the novel *piwi* family of proteins and nucleic acids. Preferred nucleic acid and amino acid sequences are described in SEQ ID NOs:1-6.

It is thus another aspect of this invention to provide a purified and isolated *piwi* family polypeptide having a role in the growth, proliferation and/or self-renewing division of stem cells, and proliferation of primordial germ cells.

The foregoing aspects and embodiments have broad utility given the biological significance of stem cell growth, proliferation and renewal, and primordial germ cell proliferation. By way of example, the foregoing aspects and embodiments are useful in the preparation of screening assays and assay kits that are used to identify compounds that affect or modulate *piwi* family biological activity, or that are used to detect the presence of the proteins and

-9-

nucleic acids of this invention in biological samples. Additionally, it is well known that isolated and purified polypeptides have utility as feed additives for livestock and further polynucleotides encoding the polypeptides are thus useful in producing the polypeptides.

5           Some of the aspects and objects of the invention having been stated hereinabove, other aspects and objects will become evident as the description proceeds, when taken in connection with the accompanying Examples and Drawings as best described hereinbelow.

#### Brief Description of the Drawings

10           Figure 1 is a molecular map of the *piwi* locus. Restriction enzyme map for 15 kb of genomic DNA surrounding the *piwi* mutations is shown. Three transcripts are represented by thick arrows indicating their relative positions, sizes, splice sites and direction of transcription. The relative insertion sites of the P-element alleles is indicated above the restriction map with orientation of  
15           insertion indicated by black (5') and white (3') boxes. The P insertions are not shown according to the scale, with *piwi*<sup>2</sup> insertion partially mapped. Restriction sites: R, *EcoRI*; H, *HindIII*; P, *PstI*; S, *Sall*; X, *XhoI*.

            Figure 2 depicts alignment of predicted PIWI protein sequence with its homologs in the *piwi* box domain in 22 different proteins. The alignments were  
20           generated using Block Maker™ analysis (Henikoff et al., 1995) against amino acids 740-782 of the predicted *piwi* sequence. The amino acid sequences were deduced from genomic DNA sequences for *C. elegans* (Ce G) or from cDNA (EST) sequences (E) for *Rattus* sp. (Rs) and *Oryza sativa* (Os). GenBank accession numbers and cosmid designations (Ce) are indicated for  
25           each sequence. Residues identical to the consensus are highlighted. Shown below the alignments, residues which are absolutely conserved are shown as capital letters and residues which are 80% or more conserved are shown as lower case letters while dashes indicate their spacing.

            Figure 3 is a graph showing that *prg-RNAi* significantly reduces germline  
30           proliferation in *C. elegans* and that reducing *prg* activity affects fertility. The graph indicates the number of progeny produced by a single F1 worm from an

-10-

RNAi *prg-1* or pBST injected mother vs. the percentage of F1 animals which produced that given number of progeny. Comparison of *prg-RNAi* (striped bars) with pBST RNAi (solid bars) animals reveals two distinctive distributions.

Figure 4 depicts the construction of full length *hiwi* cDNA clone. The 2.3kb cDNA fragment obtained from library screening (herein named 3-12) was subcloned into a pBluescript® KS<sup>+</sup> plasmid vector, the 3-12 insert is shown as an open box, the polycloning site on the vector is shown as a dark line. The 1.4kb cDNA fragment obtained from 5'RACE (herein named 1501) was subcloned into a pGEM-T Easy™ vector. The 1501 in pGEM-T Easy™ vector was digested with SacI to construct a vector for subsequent ligation. The 3-12 in pBluescript® KS<sup>+</sup> vector was partially digested with SacI to obtain the insert. This partially digested 3-12 insert DNA was used to do ligation with 1501 in pGEM-T easy vector to construct a full length *hiwi* cDNA clone.

Figure 5 is a schematic depicting the *miwi* locus (genomic location and gene structure).

Figure 5A is a schematic depicting mouse chromosome 5.

Figure 5B is a schematic depicting the *miwi* genomic locus.

Figure 5C is a schematic depicting the location and structure on the *miwi* cDNA.

Figure 5D is a schematic depicting the location and structure of *miwi* genomic clones.

Figure 6 depicts alignment of MIWI protein sequence with HIWI, *Drosophila* PIWI and *C. elegans* PRG-1 and PRG-2. Names of the sequences are shown on the right and the positions of them are shown on the left. Amino acid residues identical among above four genes are boxed in black. The high identical C-terminal domains of 104 amino acid residues are underlined.

Figure 7 is a schematic diagram of a *Drosophila* germarium. The schematized germarium is oriented with the apical end toward the left. The numbers below the germarium in the brackets denote distinct regions within the germarium. Germarial region 1 extends from the apical terminal filament cells (TFC, including cap cells, CpC) to the zone of mitotically active cysts. Germarial regions 2a and 2b include differentiating 16-cell cysts. Two somatic



stem cells (SSC) are located at the border between regions 2a and 2b. Germarial region 3 contains a newly formed egg chamber that is preparing to bud off from the germarium and contribute to the developing string of egg chambers which comprise the rest of the ovariole. Abbreviations: GSC, germline stem cell; Sp, spectrosome; ISC, inner sheath cell; Cb, cystoblast; Fu, fusome; MC, mitotically active cysts; DC, differentiating 16-cell cysts; FC, follicle cell.

Figure 8 is a schematic diagram of the myc-PIWI and PIWI-myc constructs. The *myc-piwi* transgene encodes a 97.5 kDa protein. In the myc-PIWI construct, the myc epitope is inserted after the third amino acid residue of PIWI. In the PIWI-myc construct, the epitope is inserted 13 amino acids downstream of the highly conserved PIWI box (hatched), 48 residues away from the C-terminus. Abbreviations: NH<sub>2</sub>, the amino terminus; COOH, the carboxy terminus.

Figure 9 is a graph displaying the relative division rates of wild-type and *piwi*<sup>-</sup> GSCs and demonstrating that PIWI acts cell-autonomously in GSCs to promote their division. The number of germaria counted is given in parentheses.

Figure 10 is a graph displaying a comparison of the number of spectrosomes present per wild-type germarium versus per *piwi*-overexpressed germarium and demonstrating that somatic PIWI over-expression increases the number of germline stem cells. The number of germaria examined (n) is in parentheses.

#### Detailed Description of the Invention

The present invention discloses a novel class of genes required for the growth, proliferation and self-renewing division of stem cells, and proliferation of primordial germ cells. The present invention relates to the discovery of a novel class of evolutionarily conserved genes, the *piwi* family genes, that encode a novel group of basic proteins important for stem cell division, primordial germ cell proliferation and other developmental processes.

The disclosure of the present invention demonstrates that PIWI protein can localize either in the nucleoplasm or cytoplasm and that the dose of the PIWI protein determines both the number and the mitotic rate of germline stem cells, and that the PIWI protein is localized in the germplasm (the cytoplasm of future germ cells) and plays a key role in determining the initial number of germ cells in during embryogenesis. The mouse PIWI protein (i.e., MIWI protein) is also a cytoplasmic factor with very similar functions in germ cell and stem cell determination.

To identify genes essential for stem cell division, germline stem cells in the *Drosophila* ovary were established as a model by directly identifying these cells and demonstrating their self-renewing asymmetric division pattern. This stem cell model allowed applicant to effectively circumvent difficulties intrinsic to studying mammalian stem cells, such as their rarity, lack of unique morphological markers, hyper-sensitivity to *in vitro* manipulation, and inaccessibility to systematic genetic screen. Systematic genetic screens were then conducted to identify mutations that either abolish the self-renewing ability of stem cells or the asymmetry of stem cell division, or both. These screens led to the identification of the *Drosophila piwi* gene that is essential for germline stem cell self-renewal. Subsequent cloning and molecular analysis of the *piwi* gene and its homologs in *C. elegans*, the mouse, and human, and the identification of *piwi*-like genes in *Arabidopsis* led to the following novel aspects of the present invention.

#### Definitions and Techniques Affecting

##### Gene Products and Genes

As used in the following detailed description and in the claims, the term "*piwi* family" refers to a family or group of genes and gene products including, but not limited to, PIWI, HIWI, MIWI, PRG-1 and PRG-2 proteins, and the *piwi*, *hiwi*, *miwi*, *prg-1* and *prg-2* genes, each of which are further defined herein. The term "*piwi* family" also includes other members of the *piwi* family of genes and gene products characterized by biological activity, including but not limited

to the biological activities of modulating growth, proliferation and/or self-renewing division of stem cells, and/or proliferation of primordial germ cells.

Preferably, *piwi* family genes and gene products are isolated from eukaryotic sources. Thus, the term "*piwi* family" also includes invertebrate  
5 homologs. The term "*piwi* family" further includes vertebrate homologs of *piwi* family members, including, but not limited to, mammalian, avian and fish homologs. Preferred mammalian homologs of *piwi* family members include, but are not limited to, murine and human homologs.

The terms "*piwi* family gene product", "*piwi* family protein" and "*piwi*  
10 family polypeptide" refer to proteins having amino acid sequences which are substantially identical to the native amino acid sequences in the *piwi* family and which are biologically active in that they are capable of modulating growth, proliferation and/or self-renewing division of stem cells, and/or proliferation of primordial germ cells, or cross-reacting with anti-*piwi* family antibodies raised  
15 against a *piwi* family polypeptide.

The terms "*piwi* family gene product", "*piwi* family protein" and "*piwi*  
family polypeptide" also include analogs of *piwi* family molecules which exhibit at least some biological activity in common with native *piwi* family gene products. Furthermore, those skilled in the art of mutagenesis will appreciate  
20 that other analogs, as yet undisclosed or undiscovered, may be used to construct *piwi* family analogs. There is no need for an "*piwi* family gene product", "*piwi* family protein" or "*piwi* family polypeptide" to comprise all, or substantially all of the amino acid sequence of a native *piwi* family gene product. Shorter or longer sequences are anticipated to be of use in the  
25 invention. Thus, the term "*piwi* family gene product" also includes fusion or recombinant *piwi* family polypeptides and proteins. Methods of preparing such proteins are described herein in the Examples, among other places.

The terms "*piwi* family gene", "*piwi* family gene sequence" and "*piwi*  
family gene segment" refer to any DNA sequence that is substantially identical  
30 to a polynucleotide sequence encoding a *piwi* family gene product, *piwi* family protein or *piwi* family polypeptide as defined above. The terms also refer to RNA, or antisense sequences, compatible with such DNA sequences. A "*piwi*

family gene", "*piwi* family gene sequence" and "*piwi* family gene segment" may also comprise any combination of associated control sequences.

The term "substantially identical", when used to define either a *piwi* family gene product or *piwi* family amino acid sequence, or a *piwi* family gene or *piwi* family nucleic acid sequence, means that a particular sequence, for example, a mutant sequence, varies from the sequence of a natural *piwi* family by one or more deletions, substitutions, or additions, the net effect of which is to retain at least some of biological activity of *piwi* family. Alternatively, DNA analog sequences are "substantially identical" to specific DNA sequences disclosed herein if: (a) the DNA analog sequence is derived from coding regions of the natural *piwi* family gene; or (b) the DNA analog sequence is capable of hybridization of DNA sequences of (a) under moderately stringent conditions and which encode biologically active *piwi* family gene product; or (c) the DNA sequences are degenerative as a result of the genetic code to the DNA analog sequences defined in (a) and/or (b). Substantially identical analog proteins will be greater than about 60% identical to the corresponding sequence of the native protein. Sequences having lesser degrees of similarity but comparable biological activity are considered to be equivalents. In determining nucleic acid sequences, all subject nucleic acid sequences capable of encoding substantially similar amino acid sequences are considered to be substantially similar to a reference nucleic acid sequence, regardless of differences in codon sequences.

#### Percent Similarity

Percent similarity may be determined, for example, by comparing sequence information using the GAP computer program, available from the University of Wisconsin Geneticist Computer Group. The GAP program utilizes the alignment method of Needleman et al. 1970, as revised by Smith et al. 1981. Briefly, the GAP program defines similarity as the number of aligned symbols (i.e. nucleotides or amino acids) which are similar, divided by the total number of symbols in the shorter of the two sequences. The preferred default parameters for the GAP program include: (1) a unitary comparison matrix

(containing a value of 1 for identities and 0 for non-identities) of nucleotides and the weighted comparison matrix of Gribskov et al., 1986, as described by Schwartz et al., 1979; (2) a penalty of 3.0 for each gap and an additional 0.01 penalty for each symbol and each gap; and (3) no penalty for end gaps. Other  
5 comparison techniques are described in the Examples.

The term "homology" describes a mathematically based comparison of sequence similarities which is used to identify genes or proteins with similar functions or motifs. Accordingly, the term "homology" is synonymous with the term "similarity" and "percent similarity" as defined above. Thus, the phrases  
10 "substantial homology" or "substantial similarity" have similar meanings.

### Nucleic Acid Sequences

In certain embodiments, the invention concerns the use of *piwi* family genes and gene products that include within their respective sequences a sequence which is essentially that of a *piwi* family gene, or the corresponding  
15 protein. The term "a sequence essentially as that of a *piwi* family gene", means that the sequence substantially corresponds to a portion of a *piwi* family polypeptide or *piwi* family gene and has relatively few bases or amino acids (whether DNA or protein) which are not identical to those of a *piwi* family protein or *piwi* family gene, (or a biologically functional equivalent of, when  
20 referring to proteins). The term "biologically functional equivalent" is well understood in the art and is further defined in detail herein. Accordingly, sequences which have between about 70% and about 80%; or more preferably, between about 81% and about 90%; or even more preferably, between about 91% and about 99%; of amino acids which are identical or  
25 functionally equivalent to the amino acids of a *piwi* family protein or *piwi* family gene, will be sequences which are "essentially the same".

*piwi* family gene products and *piwi* family genes which have functionally equivalent codons are also covered by the invention. The term "functionally equivalent codon" is used herein to refer to codons that encode the same  
30 amino acid, such as the six codons for arginine or serine, and also to refer to codons that encode biologically equivalent amino acids (see Table 1). Thus,

-16-

when referring to the sequence examples presented in SEQ ID NOs:1-6, the substitution of functionally equivalent codons of Table 1 into the sequence examples of SEQ ID NOs:1-6 is envisioned. Thus, applicants are in possession of amino acid and nucleic acids sequences which include such substitutions but which are not set forth herein in their entirety for convenience.

**TABLE 1 - Functionally Equivalent Codons.**

Amino Acids			Codons
Alanine	Ala	A	GCA GCC GCG GCU
Cysteine	Cys	C	UGC UGU
Aspartic Acid	Asp	D	GAC GAU
Glumatic acid	Glu	E	GAA GAG
Phenylalanine	Phe	F	UUC UUU
Glycine	Gly	G	GGA GGC GGG GGU
Histidine	His	H	CAC CAU
Isoleucine	Ile	I	AUA AUC AUU
Lysine	Lys	K	AAA AAG
Leucine	Leu	L	UUA UUG CUA CUC CUG CUU
Methionine	Met	M	AUG
Asparagine	Asn	N	AAC AAU
Proline	Pro	P	CCA CCC CCG CCU
Glutamine	Gln	Q	CAA CAG
Arginine	Arg	R	AGA AGG CGA CGC CGG CGU
Serine	Ser	S	ACG AGU UCA UCC UCG UCU
Threonine	Thr	T	ACA ACC ACG ACU
Valine	Val	V	GUA GUC GUG GUU
Tryptophan	Trp	W	UGG
Tyrosine	Tyr	Y	UAC UAU

It will also be understood that amino acid and nucleic acid sequences may include additional residues, such as additional N- or C-terminal amino acids or 5' or 3' sequences, and yet still be essentially as set forth in one of the sequences disclosed herein, so long as the sequence meets the criteria set forth above, including the maintenance of biological protein activity where protein expression is concerned. The addition of terminal sequences particularly applies to nucleic acid sequences which may, for example, include various non-coding sequences flanking either of the 5' or 3' portions of the coding region or may include various internal sequences, i.e., introns, which are known to occur within genes.

The present invention also encompasses the use of DNA segments which are complementary, or essentially complementary, to the sequences set

forth in the specification. Nucleic acid sequences which are "complementary" are those which are base-pairing according to the standard Watson-Crick complementarity rules. As used herein, the term "complementary sequences" means nucleic acid sequences which are substantially complementary, as may  
5 be assessed by the same nucleotide comparison set forth above, or as defined as being capable of hybridizing to the nucleic acid segment in question under relatively stringent conditions such as those described herein. A particular example of a contemplated complementary nucleic acid segment is an antisense oligonucleotide.

10 Nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, or organic solvents, in addition to the base composition, length of the complementary strands, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. Stringent temperature conditions  
15 will generally include temperatures in excess of 30°C, typically in excess of 37°C, and preferably in excess of 45°C. Stringent salt conditions will ordinarily be less than 1,000 mM, typically less than 500 mM, and preferably less than 200 mM. However, the combination of parameters is much more important than the measure of any single parameter. (See, e.g., Wetmur & Davidson,  
20 1968). Determining appropriate hybridization conditions to identify and/or isolate sequences containing high levels of homology is well known in the art. (See e.g., Sambrook *et al.*, 1989). For the purposes of specifying conditions of high stringency, preferred conditions are salt concentration of about 200 mM and temperature of about 45°C.

25 Probe sequences may also hybridize specifically to duplex DNA under certain conditions to form triplex or other higher order DNA complexes. The preparation of such probes and suitable hybridization conditions are well known in the art.

30 As used herein, the term "DNA segment" refers to a DNA molecule which has been isolated free of total genomic DNA of a particular species. Furthermore, a DNA segment encoding a *piwi* family polypeptide refers to a DNA segment which contains *piwi* family coding sequences, yet is isolated

away from, or purified free from, total genomic DNA of a source species, such as *Homo sapiens*. Included within the term "DNA segment" are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phages, viruses, and the like.

5           Similarly, a DNA segment comprising an isolated or purified *piwi* family gene refers to a DNA segment including *piwi* family coding sequences isolated substantially away from other naturally occurring genes or protein encoding sequences. In this respect, the term "gene" is used for simplicity to refer to a functional protein, polypeptide or peptide encoding unit. As will be understood  
10 by those in the art, this functional term includes both genomic sequences and cDNA sequences. "Isolated substantially away from other coding sequences" means that the gene of interest, in this case, the *piwi* family gene, forms the significant part of the coding region of the DNA segment, and that the DNA segment does not contain large portions of naturally-occurring coding DNA,  
15 such as large chromosomal fragments or other functional genes or cDNA coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

In particular embodiments, the invention concerns isolated DNA  
20 segments and recombinant vectors incorporating DNA sequences which encode a *piwi* family polypeptide that includes within its amino acid sequence an amino acid sequence selected from any of SEQ ID NOs:2, 4 and 6. In other particular embodiments, the invention concerns isolated DNA segments and recombinant vectors incorporating DNA sequences which encode a protein that  
25 includes within its amino acid sequence the amino acid sequence of a *piwi* family polypeptide corresponding to human tissues (referred to herein as HIWI).

It will also be understood that this invention is not limited to the particular nucleic acid and amino acid sequences of any of SEQ ID NOs:1-6. Recombinant vectors and isolated DNA segments may therefore variously  
30 include the *piwi* family polypeptide-encoding region itself, include coding regions bearing selected alterations or modifications in the basic coding region, or include encoded larger polypeptides which nevertheless include *piwi* family



polypeptide-encoding regions or may encode biologically functional equivalent proteins or peptides which have variant amino acid sequences.

In certain embodiments, the invention concerns isolated DNA segments and recombinant vectors which encode a protein or peptide that includes within  
5 its amino acid sequence an amino acid sequence essentially as set forth in any of SEQ ID NOs:2, 4 and 6. Naturally, where the DNA segment or vector encodes a full length *piwi* family gene product, the most preferred nucleic acid sequence is that which is essentially as set forth in any of SEQ ID NOs:1, 3 and 5 and which encode a protein that exhibits activity in the modulation of  
10 growth, proliferation and self-renewing division of stem cells (in for example germline stem cells), and proliferation of primordial germ cells as may be determined by, for example, immunofluorescence assays, as disclosed herein.

The term "a sequence essentially as set forth in any of SEQ ID NOs:2, 4 and 6" means that the sequence substantially corresponds to a portion an  
15 amino acid sequence of any of SEQ ID NOs:2, 4 and 6 and has relatively few amino acids which are not identical to, or a biologically functional equivalent of, the amino acids of an amino acid sequence of any of SEQ ID NOs:2, 4 and 6. The term "biologically functional equivalent" is well understood in the art and is further defined in detail herein. Accordingly, sequences, which have  
20 between about 70% and about 80%; or more preferably, between about 81% and about 90%; or even more preferably, between about 91% and about 99%; of amino acids which are identical or functionally equivalent to the amino acids of any of SEQ ID NOs:2, 4 and 6, will be sequences which "a sequence essentially as set forth in any of SEQ ID NOs:2, 4 and 6".

25 In particular embodiments, the invention concerns gene therapy methods that use isolated DNA segments and recombinant vectors incorporating DNA sequences which encode a protein that includes within its amino acid sequence an amino acid sequence of any of SEQ ID NOs:2, 4 and 6, any of SEQ ID NOs:2, 4 and 6 including sequences which are derived from  
30 mammalian tissue, among other tissue. In other particular embodiments, the invention concerns isolated DNA sequences and recombinant DNA vectors incorporating DNA sequences which encode a protein that includes within its

amino acid sequence the amino acid sequence of the *piwi* family protein from human tissue (HIWI).

In certain other embodiments, the invention concerns isolated DNA segments and recombinant vectors that include within their sequence a nucleic acid sequence essentially as set forth in any of SEQ ID NOs:1, 3 and 5. The term "a sequence essentially as set forth in any of SEQ ID NOs:1, 3 and 5" are used in the same sense as described above and means that the nucleic acid sequence substantially corresponds to a portion of any of SEQ ID NOs:1, 3 and 5, respectively, and has relatively few codons which are not identical, or functionally equivalent, to the codons of any of SEQ ID NOs:1, 3 and 5, respectively. Again, DNA segments which encode gene products exhibiting activity in the modulation of growth, proliferation and self-renewing division of stem cells, and proliferation of primordial germ cells or other biological activity of the *piwi* family gene product will be most preferred. The term "functionally equivalent codon" is used herein to refer to codons that encode the same amino acid, such as the six codons for arginine or serine, and also to refer to codons that encode biologically equivalent amino acids (see Table 1).

The nucleic acid segments of the present invention, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, enhancers, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, nucleic acid fragments may be prepared which include a short stretch complementary to a nucleic acid sequence set for in any of SEQ ID NOs:1, 3 and 5 respectively, such as about 10 nucleotides, and which are up to 10,000 or 5,000 base pairs in length, with segments of 3,000 being preferred in certain cases. DNA segments with total lengths of about 1,000, 500, 200, 100 and about 50 base pairs in length are also contemplated to be useful.

The DNA segments of the present invention encompass biologically functional equivalent *piwi* family proteins and peptides. Such sequences may rise as a consequence of codon redundancy and functional equivalency which are known to occur naturally within nucleic acid sequences and the proteins thus encoded.

Alternatively, functionally equivalent proteins or peptides may be created via the application of recombinant DNA technology, in which changes in the protein structure may be engineered, based on considerations of the properties of the amino acids being exchanged, e.g. substitution of Ile for Leu at amino acid 13, at amino acid 106 and/or at amino acid 785 for PIWI in SEQ ID NOs:1-2; substitution of Ile for Leu at amino acid 90, at amino acid 383 and/or at amino acid 816 for MIWI in SEQ ID NOs:3-4; and substitution of Ile for Leu at amino acid 76, at amino acid 303 and/or at amino acid 735 for HIWI in SEQ ID NOs:5-6. Changes designed by man may be introduced through the application of site-directed mutagenesis techniques, e.g., to introduce improvements to the antigenicity of the protein or to test *piwi* family mutants in order to examine activity in the modulation of growth, proliferation and self-renewing division of stem cells, and proliferation of primordial germ cells, or other activity at the molecular level.

If desired, one may also prepare fusion proteins and peptides, e.g., where the *piwi* family coding region is aligned within the same expression unit with other proteins or peptides having desired functions, such as for purification or immunodetection purposes (e.g., proteins which may be purified by affinity chromatography and enzyme label coding regions, respectively).

Recombinant vectors form important further aspects of the present invention. Particularly useful vectors are contemplated to be those vectors in which the coding portion of the DNA segment is positioned under the control of a promoter. The promoter may be in the form of the promoter which is naturally associated with the *piwi* family gene, e.g., in mammalian tissues, as may be obtained by isolating the 5' non-coding sequences located upstream of the coding segment or exon, for example, using recombinant cloning and/or PCR technology, in connection with the compositions disclosed herein.

In other embodiments, it is contemplated that certain advantages will be gained by positioning the coding DNA segment under the control of a recombinant, or heterologous, promoter. As used herein, a recombinant or heterologous promoter is intended to refer to a promoter that is not normally associated with a *piwi* family gene in its natural environment. Such promoters may include promoters isolated from bacterial, viral, eukaryotic, fish, avian or mammalian cells. Naturally, it will be important to employ a promoter that effectively directs the expression of the DNA segment in the cell type chosen for expression. The use of promoter and cell type combinations for protein expression is generally known to those of skill in the art of molecular biology, for example, see Sambrook et al., 1989, specifically incorporated herein by reference. The promoters employed may be constitutive, or inducible, and can be used under the appropriate conditions to direct high level expression of the introduced DNA segment, such as is advantageous in the large-scale production of recombinant proteins or peptides. Appropriate promoter systems contemplated for use in high-level expression include, but are not limited to, the vaccinia virus promoter and the baculovirus promoter.

In an alternative embodiment, the present invention provides an expression vector comprising a polynucleotide that encodes a *piwi* family polypeptide having activity in the modulation of growth, proliferation and self-renewing division of stem cells (including germline and somatic stem cells), cell proliferation (including activity in embryonic development as described in Example 1), proliferation of primordial germ cells, or other biological activity in accordance with the present invention. Also preferably, an expression vector of the present invention comprises a polynucleotide that encodes human *piwi* family gene product (HIWI). More preferably, an expression vector of the present invention comprises a polynucleotide that encodes a polypeptide comprising an amino acid residue sequence of any of SEQ ID NOs:2, 4 and 6. More preferably, an expression vector of the present invention comprises a polynucleotide comprising the nucleotide base sequence of any of SEQ ID NOs:1, 3 and 5. Even more preferably, an expression vector of the invention comprises a polynucleotide operatively linked to an enhancer-promoter. More

-23-

preferably still, an expression vector of the invention comprises a polynucleotide operatively linked to a prokaryotic promoter. Alternatively, an expression vector of the present invention comprises a polynucleotide operatively linked to an enhancer-promoter that is a eukaryotic promoter, and the expression vector further comprises a polyadenylation signal that is positioned 3' of the carboxy-terminal amino acid and within a transcriptional unit of the encoded polypeptide.

In yet another embodiment, the present invention provides a recombinant host cell transfected with a polynucleotide that encodes a *piwi* family polypeptide having activity in the modulation of growth, proliferation and self-renewing division of stem cells (including germline and somatic stem cells), cell proliferation (including activity in embryonic development as described in Example 1), proliferation of primordial germ cells, or other biological activity in accordance with the present invention. SEQ ID NOs:3-6 set forth nucleotide and amino acid sequences from representative vertebrates, human and mouse. Also contemplated by the present invention are homologous or biologically equivalent polynucleotides and *piwi* family polypeptides found in other vertebrates. Preferably, a recombinant host cell of the present invention is transfected with the polynucleotide that encodes human *piwi* family polypeptide. More preferably, a recombinant host cell of the present invention is transfected with the polynucleotide sequence of SEQ ID NO:3 (*miwi*) or SEQ ID NO:5 (*hiwi*). Even more preferably, a host cell of the invention is a eukaryotic host cell. Still more preferably, a recombinant host cell of the present invention is a vertebrate cell. Preferably, a recombinant host cell of the invention is a mammalian, avian or fish cell.

In another aspect, a recombinant host cell of the present invention is a prokaryotic host cell. Preferably, a recombinant host cell of the invention is a bacterial cell, preferably a strain of *Escherichia coli*. More preferably, a recombinant host cell comprises a polynucleotide under the transcriptional control of regulatory signals functional in the recombinant host cell, wherein the regulatory signals appropriately control expression of the *piwi* family

-24-

polypeptide in a manner to enable all necessary transcriptional and post-transcriptional modification.

In yet another embodiment, the present invention contemplates a method of preparing a *piwi* family polypeptide comprising transfecting a cell with polynucleotide that encodes a *piwi* family polypeptide having activity in the modulation of growth, self-renewing division of stem cells (including germline and somatic stem cells), cell proliferation (including activity in embryonic development as described in the Examples), proliferation of primordial germ cells or other biological activity in accordance with the present invention, to produce a transformed host cell; and maintaining the transformed host cell under biological conditions sufficient for expression of the polypeptide. More preferably, the transformed host cell is a eukaryotic cell. More preferably still, the eukaryotic cell is a vertebrate cell. Alternatively, the host cell is a prokaryotic cell. More preferably, the prokaryotic cell is a bacterial cell of *Escherichia coli*. Even more preferably, a polynucleotide transfected into the transformed cell comprises the nucleotide base sequence of any of SEQ ID NOs:1, 3 and 5. SEQ ID NOs:3-6 set forth nucleotide and amino acid sequences for representative vertebrates, mouse and human. Also contemplated by the present invention are homologues or biologically equivalent *piwi* family polynucleotides and polypeptides found in other vertebrates, particularly warm blooded vertebrates.

As mentioned above, in connection with expression embodiments to prepare recombinant *piwi* family proteins and peptides, it is contemplated that longer DNA segments will most often be used, with DNA segments encoding the entire *piwi* family protein, functional domains or cleavage products thereof, being most preferred. However, it will be appreciated that the use of shorter DNA segments to direct the expression of *piwi* family peptides or epitopic core regions, such as may be used to generate anti-*piwi* family antibodies, also falls within the scope of the invention.

DNA segments which encode peptide antigens from about 15 to about 50 amino acids in length, or more preferably, from about 15 to about 30 amino acids in length are contemplated to be particularly useful. DNA segments

encoding peptides will generally have a minimum coding length in the order of about 45 to about 150, or to about 90 nucleotides. DNA segments encoding full length proteins may have a minimum coding length on the order of about 4,000 nucleotides for a protein in accordance with any of SEQ ID NOs: 2, 4 and 6.

5 Naturally, the present invention also encompasses DNA segments which are complementary, or essentially complementary, to the sequences set forth in any of SEQ ID NOs:1, 3 and 5. The terms "complementary" and "essentially complementary" are defined above. Excepting intronic or flanking regions, and  
10 allowing for the degeneracy of the genetic code, sequences which have between about 70% and about 80%; or more preferably, between about 81% and about 90%; or even more preferably, between about 91% and about 99%; of nucleotides which are identical or functionally equivalent (i.e. encoding the same amino acid) of nucleotides of any of SEQ ID NOs:1, 3 and 5, will be  
15 sequences which are "a sequence essentially as set forth in any of SEQ ID NOs:1, 3 and 5". Sequences which are essentially the same as those set forth in any of SEQ ID NOs:1, 3 and 5 may also be functionally defined as sequences which are capable of hybridizing to a nucleic acid segment containing the complement of any of SEQ ID NOs:1, 3 and 5 under relatively  
20 stringent conditions. Suitable relatively stringent hybridization conditions are described herein and will be well known to those of skill in the art.

#### Biologically Functional Equivalents

As mentioned above, modification and changes may be made in the structure of the *piwi* family proteins and peptides described herein and still  
25 obtain a molecule having like or otherwise desirable characteristics. For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive capacity with structures such as, for example, in the nucleus of a cell. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional  
30 activity, certain amino acid sequence substitutions can be made in a protein sequence (or, of course, its underlying DNA coding sequence) and

nevertheless obtain a protein with like or even countervailing properties (e.g., antagonistic v. agonistic). It is thus contemplated by the inventor that various changes may be made in the sequence of the *piwi* family proteins and peptides (or underlying DNA) without appreciable loss of their biological utility or activity.

5 It is also well understood by the skilled artisan that, inherent in the definition of a biologically functional equivalent protein or peptide, is the concept that there is a limit to the number of changes that may be made within a defined portion of the molecule and still result in a molecule with an acceptable level of equivalent biological activity. Biologically functional  
10 equivalent peptides are thus defined herein as those peptides in which certain, not most or all, of the amino acids may be substituted. Of course, a plurality of distinct proteins/peptides with different substitutions may easily be made and used in accordance with the invention.

15 It is also well understood that where certain residues are shown to be particularly important to the biological or structural properties of a protein or peptide, e.g., residues in active sites, such residues may not generally be exchanged. This is the case in the present invention, where if any changes, for example, in the PIWI box of *piwi* family polypeptides, could result in a loss of an aspect of the utility of the resulting peptide for the present invention.

20 Amino acid substitutions, such as those which might be employed in modifying the *piwi* family proteins and peptides described herein, are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. An analysis of the size, shape and type of the amino acid side-chain substituents  
25 reveals that arginine, lysine and histidine are all positively charged residues; that alanine, glycine and serine are all a similar size; and that phenylalanine, tryptophan and tyrosine all have a generally similar shape. Therefore, based upon these considerations, arginine, lysine and histidine; alanine, glycine and serine; and phenylalanine, tryptophan and tyrosine; are defined herein as  
30 biologically functional equivalents.

In making such changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the



basis of their hydrophobicity and charge characteristics, these are: isoleucine (+ 4.5); valine (+ 4.2); leucine (+ 3.8); phenylalanine (+ 2.8); cysteine/cystine (+ 2.5); methionine (+ 1.9); alanine (+ 1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2);  
5 glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte & Doolittle, 1982, incorporated herein by reference). It is known that  
10 certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  of the original value is preferred, those which are within  $\pm 1$  of the original value are particularly preferred, and those  
15 within  $\pm 0.5$  of the original value are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Pat. No. 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent  
20 amino acids, correlates with its immunogenicity and antigenicity, i.e. with a biological property of the protein. It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent protein.

As detailed in U.S. Pat. No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+ 3.0); lysine (+ 3.0); aspartate (+ 3.0 $\pm$ 1); glutamate (+ 3.0 $\pm$ 1); serine (+ 0.3); asparagine (+ 0.2); glutamine (+ 0.2); glycine (0); threonine (-0.4); proline (-0.5 $\pm$ 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4).  
25

30 In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  of the original value is preferred, those which are within  $\pm 1$  of the original value are

-28-

particularly preferred, and those within  $\pm 0.5$  of the original value are even more particularly preferred.

While discussion has focused on functionally equivalent polypeptides arising from amino acid changes, it will be appreciated that these changes may be effected by alteration of the encoding DNA, taking into consideration also that the genetic code is degenerate and that two or more codons may code for the same amino acid.

#### Sequence Modification Techniques

Modifications to the *piwi* family proteins and peptides described herein may be carried out using techniques such as site directed mutagenesis. Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent proteins or peptides, through specific mutagenesis of the underlying DNA. The technique further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 30 nucleotides in length is preferred, with about 5 to 10 residues on both sides of the junction of the sequence being altered.

In general, the technique of site-specific mutagenesis is well known in the art as exemplified by publications (e.g., Adelman et al., 1983). As will be appreciated, the technique typically employs a phage vector which exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage (Messing et al., 1981). These phage are readily commercially available and their use is generally well known to those skilled in the art. Double stranded plasmids are

also routinely employed in site directed mutagenesis which eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart the two strands of a double stranded vector which includes within its sequence a DNA sequence which encodes, for example, a human *piwi* family polypeptide (HIWI). An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically, for example by the method of Crea et al. (1978). This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected gene using site-directed mutagenesis is provided as a means of producing potentially useful *piwi* family polypeptide or other species capable of modulating growth, proliferation and self-renewing division of stem cells, and proliferation of primordial germ cells and is not meant to be limiting as there are other ways in which sequence variants of these peptides may be obtained. For example, recombinant vectors encoding the desired genes may be treated with mutagenic agents to obtain sequence variants (see, e.g., a method described by Eichenlaub, 1979) for the mutagenesis of plasmid DNA using hydroxylamine.

#### Other Structural Equivalents

In addition to the *piwi* family peptidyl compounds described herein, the inventors also contemplate that other sterically similar compounds may be formulated to mimic the key portions of the peptide structure. Such compounds may be used in the same manner as the peptides of the invention and hence are also functional equivalents. The generation of a structural functional

-30-

equivalent may be achieved by the techniques of modeling and chemical design known to those of skill in the art. It will be understood that all such sterically similar constructs fall within the scope of the present invention.

### Introduction of Gene Products

5           Where the gene itself is employed to introduce the gene products, a convenient method of introduction will be through the use of a recombinant vector which incorporates the desired gene, together with its associated control sequences. The preparation of recombinant vectors is well known to those of skill in the art and described in many references, such as, for example,  
10       Sambrook et al. (1989), specifically incorporated herein by reference.

          In vectors, it is understood that the DNA coding sequences to be expressed, in this case those encoding the *piwi* family gene products, are positioned adjacent to and under the control of a promoter. It is understood in the art that to bring a coding sequence under the control of such a promoter,  
15       one generally positions the 5' end of the transcription initiation site of the transcriptional reading frame of the gene product to be expressed between about 1 and about 50 nucleotides "downstream" of (i.e., 3' of) the chosen promoter. One may also desire to incorporate into the transcriptional unit of the vector an appropriate polyadenylation site (e.g., 5'-AATAAA-3'), if one was not  
20       contained within the original inserted DNA. Typically, these poly A addition sites are placed about 30 to 2000 nucleotides "downstream" of the coding sequence at a position prior to transcription termination.

          While use of the control sequences of the specific gene (i.e., a *piwi* family promoter for a *piwi* family gene) will be preferred, there is no reason why  
25       other control sequences could not be employed, so long as they are compatible with the genotype of the cell being treated. Thus, one may mention other useful promoters by way of example, including, e.g., an SV40 early promoter, a long terminal repeat promoter from retrovirus, an actin promoter, a heat shock promoter, a metallothionein promoter, and the like.

30       As is known in the art, a promoter is a region of a DNA molecule typically within about 100 nucleotide pairs in front of (upstream of) the point at

which transcription begins (i.e., a transcription start site). That region typically contains several types of DNA sequence elements that are located in similar relative positions in different genes. As used herein, the term "promoter" includes what is referred to in the art as an upstream promoter region, a promoter region or a promoter of a generalized eukaryotic RNA Polymerase II transcription unit.

Another type of discrete transcription regulatory sequence element is an enhancer. An enhancer provides specificity of time, location and expression level for a particular encoding region (e.g., gene). A major function of an enhancer is to increase the level of transcription of a coding sequence in a cell that contains one or more transcription factors that bind to that enhancer. Unlike a promoter, an enhancer can function when located at variable distances from transcription start sites so long as a promoter is present.

As used herein, the phrase "enhancer-promoter" means a composite unit that contains both enhancer and promoter elements. An enhancer-promoter is operatively linked to a coding sequence that encodes at least one gene product. As used herein, the phrase "operatively linked" means that an enhancer-promoter is connected to a coding sequence in such a way that the transcription of that coding sequence is controlled and regulated by that enhancer-promoter. Means for operatively linking an enhancer-promoter to a coding sequence are well known in the art. As is also well known in the art, the precise orientation and location relative to a coding sequence whose transcription is controlled, is dependent *inter alia* upon the specific nature of the enhancer-promoter. Thus, a TATA box minimal promoter is typically located from about 25 to about 30 base pairs upstream of a transcription initiation site and an upstream promoter element is typically located from about 100 to about 200 base pairs upstream of a transcription initiation site. In contrast, an enhancer can be located downstream from the initiation site and can be at a considerable distance from that site.

An enhancer-promoter used in a vector construct of the present invention can be any enhancer-promoter that drives expression in a cell to be

transfected. By employing an enhancer-promoter with well-known properties, the level and pattern of gene product expression can be optimized.

For introduction of, for example, the human *piwi* family gene (*hiwi*), it is proposed that one will desire to preferably employ a vector construct that will deliver the desired gene to the affected cells. This will, of course, generally require that the construct be delivered to the targeted cells, for example, mammalian, avian or fish germline stem cells. It is proposed that this may be achieved most preferably by introduction of the desired gene through the use of a viral vector to carry the *piwi* family sequence to efficiently infect the cells. These vectors will preferably be an adenoviral, a retroviral, a vaccinia viral vector or adeno-associated virus. These vectors are preferred because they have been successfully used to deliver desired sequences to cells and tend to have a high infection efficiency. Suitable vector-*piwi* family gene constructs are adapted for administration as pharmaceutical compositions, as described herein below.

Commonly used viral promoters for expression vectors are derived from polyoma, cytomegalovirus, Adenovirus 2, and Simian Virus 40 (SV40). The early and late promoters of SV40 virus are particularly useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication. Smaller or larger SV40 fragments may also be used, provided there is included the approximately 250 bp sequence extending from the *Hind* III site toward the *Bgl* I site located in the viral origin of replication. Further, it is also possible, and often desirable, to utilize promoter or control sequences normally associated with the desired gene sequence, provided such control sequences are compatible with the host cell systems.

The origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (e.g., Polyoma, Adeno, VSV, BPV) source, or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter is often sufficient.

Where a *piwi* family gene itself is employed it will be most convenient to simply use a wild type *piwi* family gene directly. However, it is contemplated

-33-

that certain regions of a *piwi* family gene may be employed exclusively without employing an entire wild type *piwi* family gene. It is proposed that it will ultimately be preferable to employ the smallest region needed to modulate stem cell growth, proliferation and self-renewing division, and/or primordial germ cell proliferation, so that one is not introducing unnecessary DNA into cells which receive a *piwi* family gene construct. Techniques well known to those of skill in the art, such as the use of restriction enzymes, will allow for the generation of small regions of an exemplary *piwi* family gene. The ability of these regions to modulate cell signaling can easily be determined by the assays reported in the Examples. In general, techniques for assessing the modulation of stem cell growth, proliferation and self-renewing division, along with primordial germ cell proliferation are known in the art.

#### Cell Culture Techniques

As disclosed herein, PIWI functions both as a cytoplasmic protein (during embryonic germline development) and as nuclear protein during stem cell division. PIWI has a cell-autonomous function to promote stem cell division in addition to its signaling function in maintaining stem cells. The dose of PIWI gene determines the number of germline stem cells formed in the embryo.

In accordance with the present invention, *piwi* family genes and gene products can thus be used to expand a primitive cell population *in vitro*. The term "primitive cell" is adopted herein for convenience to collectively refer to stem cells, primordial germ cells and to other primitive cells having totipotent or pluripotent characteristics, or other characteristics of stem cells and primordial germ cells. The term "stem cell" is meant to have its art-recognized meaning, and to include, but not be limited to, the following examples of stem cells: hematopoietic stem cells, neural stem cells, and germline stem cells. The term "primordial germ cell" is also meant to have its art-recognized meaning. Such expanded primitive cell populations are utilized for therapeutic or genetic engineering purposes in human and mammalian, avian, and fish systems.

The term "sustained" as used herein with respect to primitive cells, and cultures of the same, refers to a cell or cell culture capable of undergoing further cell division, even if the cells are eventually subject to senescence.

5 A method of culturing primitive cells is thus envisioned in accordance with the present invention. The method comprises providing a culture comprising a primitive cell, and delivering to the provided cell a *piwi* family gene product, whereby growth of the cell in the culture is accomplished. Preferably, the primitive cell is maintained in culture for a time sufficient to produce a sustained cell culture.

10 Optionally, the *piwi* family gene product is delivered to the primitive cell by delivering to the primitive cell an effective amount of a DNA molecule comprising a polynucleotide that encodes a biologically active *piwi* family polypeptide; and maintaining the primitive cell under conditions sufficient for expression of the polypeptide. Alternatively, the *piwi* family gene product is  
15 delivered to the cell by administration of a biologically active *piwi* family polypeptide to the culture via the culture media or via a feeder matrix included in the culture.

A preferred embodiment of the method comprises: (a) collecting primitive cells; (b) depositing the primitive cells in contact with a feeder matrix,  
20 the feeder matrix comprising a *piwi* family gene product; and (c) growing the primitive cells on the feeder matrix in the presence of media for a time sufficient to produce a sustained primitive cell culture.

Thus, typically, the cultures of the present invention include a feeder matrix as defined herein. A feeder matrix can either be cells or cell lines  
25 cultured for the purpose of obtaining a sustained primitive cell culture. In this case the cell line can be transformed with a *piwi* family gene product in accordance with techniques disclosed herein so as to produce the *piwi* family gene product. Alternatively, a feeder matrix can be derived from or provided by the organ or tissue in which the primitive cells are located, e.g., the gonad.  
30 Alternatively, the feeder cells comprising the feeder matrix could be substituted with extracellular matrix plus bound growth factors, and particularly including a *piwi* family gene product.



A feeder matrix as used herein is constructed in accordance with procedures known in the art. As noted above, it is preferred that the feeder matrix be preconditioned. By the term "preconditioned" it is meant that the feeder matrix is cultured in the presence of media for a period of time prior to the depositing of primitive cells in contact with the feeder matrix, e.g. a time sufficient to initiate and establish production of, for example, growth factors or other factors by the feeder matrix, and particularly including a *piwi* family gene product. Thus, a feeder matrix can be preconditioned by culturing the feeder matrix by itself for one to two days prior to the depositing of primitive cells in contact with the feeder matrix.

Candidate cells for use in the feeder matrix are disclosed in the Examples. Additionally, the feeder matrix can comprise Sertoli cells or fibroblast cells, including mouse fibroblast cells. STO fibroblasts and primary fibroblasts are suitable examples. Also, it is contemplated that feeder matrices comprising cells from other murine species (e.g., rat); other mammalian species (e.g., ungulate, bovine, and porcine species); or avian species (e.g. *Gallinacea*, chicken, turkey, duck, goose, quail and pheasant) may also be used.

The culturing method of the present invention can optionally comprise establishing a monolayer of feeder cells. Feeder cells may mitotically inactivated using standard techniques. For example, the feeder cells may be exposed to gamma radiation (e.g. 4000 Rads of gamma radiation) or may be treated with Mitomycin C (e.g. 10  $\mu$ g/ml for 2-3 hours). Procedures for mitotically inactivating cells are also detailed in the information typically sent with cells from the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209 (e.g. STO feeder cells are available under ATCC accession number 1503). Monolayers may optionally be cultured to about 80% confluency, preferably to about 90% confluency, and more preferably about 100% confluency. While configuration of the feeder cells as a monolayer is the preferred configuration for the culture, any suitable configuration is contemplated to be within the scope of the present invention. Thus, for example, layers, monolayers, clusters, aggregates or other

associations or groupings of feeder cells are contemplated to fall within the scope of the present invention and are particularly contemplated to fall with the meaning of the term "matrix".

5 Thus, the primitive cells can extend non-uniformly through the feeder cells vertically, horizontally, diagonally, or in any combination thereof, such that cell aggregates are formed. In accordance with the present invention, it is thus contemplated that the cultures can be formed by mixing primitive cells with appropriate feeder cells and inserting the mixture into a suitable culture container, such as a multi-well plate.

10 The media used in carrying out the present invention may be any suitable media. The media may be a conditioned media or a synthetic media, both of which are known in the art. Conditioned media, and particularly BRL conditioned media, is currently preferred. By way of example, BRL conditioned media is prepared according to art-recognized techniques, such as described  
15 by Smith, A. G. and Hooper, M. L., *Dev. Biol.* 1987 May; 121(1): 1-9. BRL cells are available from ATCC under accession number CRL-1442. The media can be supplemented with a *piwi* family gene product. The media can also be supplemented with growth factors, including but not limited to leukemia inhibitory factor (LIF), insulin-like growth factor (IGF), fibroblast growth factor (FGF), basic fibroblast growth factor (bFGF), stem cell factor (SCF - also called  
20 steel factor or SF), transforming growth factor- $\beta$ 1 (TGF- $\beta$ 1) and anti-retinoic acid.

Cell cultures of the present invention may be formulated for administration to animals by dissociating the cells (e.g., by mechanical  
25 dissociation) and intimately admixing the cells with a pharmaceutically acceptable carrier (e.g., phosphate-buffered saline (PBS) solution). Primitive cells in such formulations may be prepared to carry a heterologous DNA sequence into a subject, preferably a vertebrate subject and more preferably a warm-blooded vertebrate subject, in the manner described in greater detail  
30 below.

The primitive cells cultured in accordance with the present invention are useful, among other things, as a tool for the study of embryological

-37-

development (i.e., by labeling the cells with a marker gene and observing their distribution after injection *in vivo*) and the production of transgenic animals. They are useful in allowing the application of homologous recombination to the production of transgenic animals.

5           The primitive cells should also be useful for the production of chimeric animals, including transgenic chimeric animals. It is contemplated that chimeric or transgenic chimeric non-human animals, including particularly chimeric or transgenic chimeric avians (i.e. birds), will be useful for the recovery of heterologous proteins from the animals. In the case of an avian, the protein  
10       can be recovered directly from the eggs of such chimeric or transgenic chimeric avians. Such avians can be thus be used for the production and recovery of therapeutic proteins and other polypeptides, including *piwi* family polypeptides themselves.

          The following U.S. Patents pertain to the culturing of stem cells an are  
15       used herein to illuminate the background of the invention, and in particular cases, to provide additional details respecting the practice. Each of the following U.S. Patent is incorporated herein by reference in its entirety: U.S. Patent No. 5,690,926 issued November 25, 1997 to Hogan; U.S. Patent No. 5,670,372 issued September 23, 1997 to Hogan; U.S. Patent No. 5,537,357  
20       issued September 26, 1995 to Hogan; U.S. Patent No. 5,340,740 issued August 23, 1994 to Petitte et al.; U.S. Patent No. 5,656,479 issued August 12, 1997 to Petitte et al.; and U.S. Patent No. 5,830,510 issued November 3, 1998 to Petitte et al.

#### Transgenic Animals

25           It is also contemplated to be within the scope of the present invention to prepare a transgenic non-human animal which expresses a *piwi* family gene of the present invention. A preferred transgenic animal is a mouse.

          The term "transgene" refers to exogenous genetic material which does not naturally form part of the genetic material of an animal to be genetically  
30       altered but can be incorporated into the germ and/or somatic cells of that animal by standard transgenic techniques. The term "heterologous DNA"

refers to DNA which has been transferred from one individual animal, species or breed to a different individual animal, species or breed. The term "transgenic" refers to cells, tissues, embryos, fetuses or animals which carry one or more transgenes. The term "chimeric" refers to an embryo, fetus or animal which consists of two or more tissues of different genetic composition.

Techniques for the preparation of transgenic animals are known in the art. Exemplary techniques are described in U.S. Patent No. 5,489,742 (transgenic rats); U.S. Patent Nos. 4,736,866, 5,550,316, 5,614,396, 5,625,125 and 5,648,061 (transgenic mice); U.S. Patent No. 5,573,933 (transgenic pigs); 5,162,215 (transgenic avian species) and U.S. Patent No. 5,741,957 (transgenic bovine species), the entire contents of each of which are herein incorporated by reference.

With respect to a representative method for the preparation of a transgenic mouse, cloned recombinant or synthetic DNA sequences or DNA segments encoding a *piwi* family gene product are injected into fertilized mouse eggs (e.g. an embryo). The injected eggs are implanted in pseudo pregnant females and are grown to term to provide transgenic mice whose cells express a *piwi* family gene product. Preferably, the injected sequences are constructed having promoter sequences connected so as to express the desired protein in germline stem cells or other primitive cells of the transgenic mouse.

Fish represent a category of animals of interest for agricultural and ecological reasons. Representative fish species include, but are not limited to, trout, salmon, carp, shark, ray, flounder, sole, tilapia, medaka, goldfish, guppy, molly, platyfish, swordtail, zebrafish, loach, catfish, and the like. Representative methods for the preparation of a transgenic fish have been described by Ozato et al, *Cell Differ.*, 19:237-244 (1986), Inoue et al, *Cell Differ. Dev.*, 29:123-128 (1990), Rokkones et al, *J. Comp. Physiol. B*, 158:751-758 (1989), and Guyomard et al, *Biochimie*, 71:857-863 (1989), describing preparation of transgenic medaka, medaka, salmon and trout, respectively.

The term "avian" as used herein refers to any avian species, including but not limited to *Gallinacea sp.*, chicken, turkey, duck, goose, quail and pheasant. Chicken is currently preferred.

5 With respect to an exemplary method for the preparation of a transgenic avian species, certain donor cell types have been isolated that retain viability when injected into recipient embryos. See Etches et al., *Poultry Science* 72:882-887 (1993); Etches et al., in *Avian Incubation*, Chapter 22, Butterworth Publishers (1990); Verrinder Gibbins et al., *Fourth World Congress on Genetics Applied to Livestock Production*, Edinburgh, (1990); Petite et al., *Development* 108, 185-189 (1990)). These studies showed that blastodermal cells derived from Stage X embryos (embryo at oviposition) remained viable when transferred to comparable recipient Stage X embryos.

15 Thus, the present invention provides a new method of altering the phenotype of an animal, such as a bird, and the animals so produced with the primitive cells cultured in accordance with techniques disclosed herein. The method comprises transfecting primitive cells cultured in accordance with the methods disclosed herein with the DNA sequence *in vitro* (e.g., by electroporation or transformation with a retroviral vector), and then injecting the transfected primitive cells into an embryo, (e.g. an egg containing an embryonic 20 bird via the yolk sac or onto the chorioallantoic membrane), with the DNA sequence being effective to cause a change in phenotype in the embryonic animal (e.g., a change in protein expression, a change in growth rate, feed efficiency, disease resistance, or a combination of all of these factors). Preferably, the primitive cells are also transfected with a *piwi* family gene to 25 facilitate production of germ cells in the transgenic or chimeric bird and thereby also facilitate germline transmission of the DNA sequence of interest.

Preferably, in the case of a bird, the egg into which the DNA is introduced is incubated to hatch, and the bird so produced is raised to at least an age at which the change in phenotype is expressed. It is of no deleterious 30 consequence if the transformed embryo and bird is chimeric, so long as a physiological response is achieved in the animal after hatch sufficient to evoke a measurable phenotypic change measured in any suitable way. By way of

additional example, protein expression in the chimera, or transgenic offspring thereof produced by germ line transmission of the DNA sequence, may be established through incorporation of an inducible promoter into the introduced DNA sequence coupled with treatment of the chimera, or transgenic offspring thereof, with an agent which induces the promoter.

The mechanism of *in ovo* injection is not critical, but it is preferred that the method not unduly damage the tissues and organs of the embryo or the extraembryonic membranes surrounding it so that the treatment will not decrease hatch rate. For example, *in ovo* injections can be carried out by forming a window in a bird egg and then using a micro-pipette or needle for the injection, such a 50  $\mu$ m-diameter micro-pipette or a 50  $\mu$ m-diameter needle. If desired, the bird egg can be sealed with a substantially bacteria-impermeable sealing material such as wax or the like to prevent subsequent entry of undesirable bacteria. Suitable sealing material may optionally comprise biologically acceptable antimicrobial compositions.

The DNA sequence introduced with the primitive cells in accordance with the present invention is, in general, a construct comprised of a promoter functional in the cells of the candidate animal and a gene encoding a peptide or protein operably linked to the promoter. Preferably, the protein or peptide is physiologically active and capable of producing a phenotypic change in the animal. In general, the DNA construct may be a linear DNA sequence (introduced into the primitive cells of the invention by electroporation) or a sequence carried by a vector or other suitable carrier for transforming the primitive cells of the invention using various techniques known in the art, such as liposomes, calcium phosphate, electroporation, or DMSO. Vectors, as discussed herein, may be plasmids, viruses (including retroviruses), and phage, whether in native form or derivatives thereof.

The production of cloned genes, recombinant DNA, vectors, transformed host cells, proteins and protein fragments by genetic engineering is well known. See e.g., U.S. Pat. No. 4,761,371 to Bell et al. at Col. 6 line 3 to Col. 9 line 65; U.S. Pat. No. 4,877,729 to Clark et al. at Col. 4 line 38 to Col. 7 line 6; U.S. Pat. No. 4,912,038 to Schilling at Col. 3 line 26 to Col. 14 line 12. Protocols for

restriction endonuclease digestion, preparation of vectors, DNA purification and other such procedures are essentially as described in standard cloning manuals. See Sambrook et al., *Molecular Cloning, a Laboratory Manual*, (2d Ed., Cold Spring Harbor Press, N.Y. (1989)).

5           As disclosed herein, a vector is a replicable DNA construct used herein to either amplify and/or express DNA encoding the gene of interest. A suitable expression vector will have controlling elements capable of expressing the cloned cDNA or genomic DNA placed in the correct orientation when the vector is introduced into the correct host. Such elements typically include but are not  
10       limited to a promoter region which interacts specifically with cellular proteins involved in transcription or which may be inducible by administration of an inducing agent, enhancer elements which can stimulate transcription many-fold from linked heterologous promoters, a splice acceptor and/or donor sequence, and termination and polyadenylation signals. Also required is the sequence for  
15       a ribosome binding site capable of permitting translation and which is operably linked to the gene to be expressed. Recently, a muscle-specific promoter has been isolated which is positioned upstream of both the skeletal muscle structural gene and the essential proximal promoter element and is operably associated with each. (Mar and Ordahl, *Proc. Natl. Acad. Sci. USA* 85,  
20       6404-6408 (1988)). Vectors comprise plasmids, viruses (e.g. adenovirus, cytomegalovirus), phage, and DNA fragments integratable into the host genome by recombination. The vector may optionally replicate and function independently of the host genome, or may in some instances integrate into the genome itself.

25           The *piwi* family genes of the present invention can thus be used to create transgenic animals with increased number of primordial germ cells. Such animals in mammalian, avian, and fish systems should provide an improved platform for germline-mediated gene transfer to create transgenic stocks.

30           The genetically modified animals could express transgenes of commercial interest, including those having therapeutic or prophylactic value to the animal itself or to its offspring. For example, the transgene of interest

-42-

can encode a protein or peptide including, but not limited to, growth hormone, thyroid releasing hormone (TRH), Marek's MDX, and immunogenic recombinant antigens such as that for coccidiosis.

Alternatively, the genetically modified animals can be employed in the production of therapeutic agents. For example, genetically modified avian species that can lay eggs containing drugs, proteins and antibodies to ward off illness that have been produced via the insertion the genes that make the proteins into viral and other vectors to get them into the birds. *piwi* family gene products can also be produced and isolated in this manner.

In accordance with the present invention, *piwi* family genes can also be inserted into such genetically modified animals to create animals with increased number of primordial germ cells, thereby enhancing the likelihood of germline transmission of the transgene encoding the therapeutic agent due to the increased number of primordial germ cells.

#### Pharmaceutical Compositions

In a preferred embodiment, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide of the present invention and a physiologically acceptable carrier. More preferably, a pharmaceutical composition comprises a *piwi* family polypeptide or a polynucleotide that encodes those polypeptides.

A composition of the present invention is typically administered parenterally in dosage unit formulations containing standard, well-known nontoxic physiologically acceptable carriers, adjuvants, and vehicles as desired. The term "parenteral" as used herein includes intravenous, intra-muscular, intra-arterial injection, or infusion techniques.

Injectable preparations, for example sterile injectable aqueous or oleaginous suspensions, are formulated according to the known art using suitable dispersing or wetting agents and suspending agents. The sterile injectable preparation can also be a sterile injectable solution or suspension in a nontoxic parenterally acceptable diluent or solvent, for example, as a solution in 1,3-butanediol.



-43-

Among the acceptable vehicles and solvents that may be employed are water, Ringer's solution, and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose any bland fixed oil can be employed including synthetic mono- or di-glycerides. In addition, fatty acids such as oleic acid find use in the preparation of injectables.

Preferred carriers include neutral saline solutions buffered with phosphate, lactate, Tris, and the like. Of course, one purifies the vector sufficiently to render it essentially free of undesirable contaminants, such as defective interfering adenovirus particles or endotoxins and other pyrogens such that it does not cause any untoward reactions in the individual receiving the vector construct. A preferred means of purifying the vector involves the use of buoyant density gradients, such as cesium chloride gradient centrifugation.

A transfected cell can also serve as a carrier. By way of example, a liver cell can be removed from an organism, transfected with a polynucleotide of the present invention using methods set forth above and then the transfected cell returned to the organism (e.g. injected intra-vascularly).

#### Generation of Antibodies

In still another embodiment, the present invention provides an antibody immunoreactive with a polypeptide of the present invention. Preferably, an antibody of the invention is a monoclonal antibody. Means for preparing and characterizing antibodies are well known in the art (See, e.g., *Antibodies A Laboratory Manual*, E. Howell and D. Lane, Cold Spring Harbor Laboratory, 1988).

Briefly, a polyclonal antibody is prepared by immunizing an animal with an immunogen comprising a polypeptide or polynucleotide of the present invention, and collecting antisera from that immunized animal. A wide range of animal species can be used for the production of antisera. Typically an animal used for production of anti-antisera is a rabbit, a mouse, a rat, a hamster or a guinea pig. Because of the relatively large blood volume of rabbits, a rabbit is a preferred choice for production of polyclonal antibodies.

As is well known in the art, a given polypeptide or polynucleotide may vary in its immunogenicity. It is often necessary therefore to couple the immunogen (e.g., a polypeptide or polynucleotide) of the present invention) with a carrier. Exemplary and preferred carriers are keyhole limpet hemocyanin (KLH) and bovine serum albumin (BSA). Other albumins such as ovalbumin, mouse serum albumin or rabbit serum albumin can also be used as carriers.

Means for conjugating a polypeptide or a polynucleotide to a carrier protein are well known in the art and include glutaraldehyde, m-maleimidobencoyl-N-hydroxysuccinimide ester, carbodiimide and bis-biaztotized benzidine.

As is also well known in the art, immunogenicity to a particular immunogen can be enhanced by the use of non-specific stimulators of the immune response known as adjuvants. Exemplary and preferred adjuvants include complete Freund's adjuvant, incomplete Freund's adjuvants and aluminum hydroxide adjuvant.

The amount of immunogen used of the production of polyclonal antibodies varies, *inter alia*, upon the nature of the immunogen as well as the animal used for immunization. A variety of routes can be used to administer the immunogen, e.g. subcutaneous, intramuscular, intradermal, intravenous and intraperitoneal. The production of polyclonal antibodies is monitored by sampling blood of the immunized animal at various points following immunization. When a desired level of immunogenicity is obtained, the immunized animal can be bled and the serum isolated and stored.

In another aspect, the present invention contemplates a method of producing an antibody immunoreactive with a *piwi* family polypeptide, the method comprising the steps of (a) transfecting recombinant host cells with a polynucleotide that encodes that polypeptide; (b) culturing the host cells under conditions sufficient for expression of the polypeptide; (c) recovering the polypeptide; and (d) preparing antibodies to the polypeptide. Preferably, the *piwi* family polypeptide possess a biological activity in accordance with the

-45-

present invention. Even more preferably, the present invention provides antibodies prepared according to the method described above.

A monoclonal antibody of the present invention can be readily prepared through use of well-known techniques such as those exemplified in U.S. Pat. No 4,196,265, herein incorporated by reference. Typically, a technique involves first immunizing a suitable animal with a selected antigen (e.g., a polypeptide or polynucleotide of the present invention) in a manner sufficient to provide an immune response. Rodents such as mice and rats are preferred animals. Spleen cells from the immunized animal are then fused with cells of an immortal myeloma cell. Where the immunized animal is a mouse, a preferred myeloma cell is a murine NS-1 myeloma cell.

The fused spleen/myeloma cells are cultured in a selective medium to select fused spleen/myeloma cells from the parental cells. Fused cells are separated from the mixture of non-fused parental cells, for example, by the addition of agents that block the *de novo* synthesis of nucleotides in the tissue culture media. Exemplary and preferred agents are aminopterin, methotrexate, and azaserine. Aminopterin and methotrexate block *de novo* synthesis of both purines and pyrimidines, whereas azaserine blocks only purine synthesis. Where aminopterin or methotrexate is used, the media is supplemented with hypoxanthine and thymidine as a source of nucleotides. Where azaserine is used, the media is supplemented with hypoxanthine.

This culturing provides a population of hybridomas from which specific hybridomas are selected. Typically, selection of hybridomas is performed by culturing the cells by single-clone dilution in microtiter plates, followed by testing the individual clonal supernatants for reactivity with an antigen-polypeptides. The selected clones can then be propagated indefinitely to provide the monoclonal antibody.

By way of specific example, to produce an antibody of the present invention, mice are injected intraperitoneally with between about 1-200  $\mu$ g of an antigen comprising a polypeptide of the present invention. B lymphocyte cells are stimulated to grow by injecting the antigen in association with an adjuvant such as complete Freund's adjuvant (a non-specific stimulator of the

immune response containing killed *Mycobacterium tuberculosis*). At some time (e.g., at least two weeks) after the first injection, mice are boosted by injection with a second dose of the antigen mixed with incomplete Freund's adjuvant.

5 A few weeks after the second injection, mice are tail bled and the sera titrated by immunoprecipitation against radiolabeled antigen. Preferably, the method of boosting and titering is repeated until a suitable titer is achieved. The spleen of the mouse with the highest titer is removed and the spleen lymphocytes are obtained by homogenizing the spleen with a syringe. Typically, a spleen from an immunized mouse contains approximately  $5 \times 10^7$  to  
10  $2 \times 10^8$  lymphocytes.

Mutant lymphocyte cells known as myeloma cells are obtained from laboratory animals in which such cells have been induced to grow by a variety of well-known methods. Myeloma cells lack the salvage pathway of nucleotide biosynthesis. Because myeloma cells are tumor cells, they can be propagated  
15 indefinitely in tissue culture, and are thus denominated immortal. Numerous cultured cell lines of myeloma cells from mice and rats, such as murine NS-1 myeloma cells, have been established.

Myeloma cells are combined under conditions appropriate to foster fusion with the normal antibody-producing cells from the spleen of the mouse  
20 or rat injected with the antigen/polypeptide of the present invention. Fusion conditions include, for example, the presence of polyethylene glycol. The resulting fused cells are hybridoma cells. Like myeloma cells, hybridoma cells grow indefinitely in culture.

Hybridoma cells are separated from unfused myeloma cells by culturing  
25 in a selection medium such as HAT media (hypoxanthine, aminopterin, thymidine). Unfused myeloma cells lack the enzymes necessary to synthesize nucleotides from the salvage pathway because they are killed in the presence of aminopterin, methotrexate, or azaserine. Unfused lymphocytes also do not continue to grow in tissue culture. Thus, only cells that have successfully fused  
30 (hybridoma cells) can grow in the selection media.

Each of the surviving hybridoma cells produces a single antibody. These cells are then screened for the production of the specific antibody

-47-

immunoreactive with an antigen/polypeptide of the present invention. Single cell hybridomas are isolated by limiting dilutions of the hybridomas. The hybridomas are serially diluted many times and, after the dilutions are allowed to grow, the supernatant is tested for the presence of the monoclonal antibody.

5 The clones producing that antibody are then cultured in large amounts to produce an antibody of the present invention in convenient quantity.

By use of a monoclonal antibody of the present invention, specific polypeptides and polynucleotide of the invention can be recognized as antigens, and thus identified. Once identified, those polypeptides and polynucleotide can be isolated and purified by techniques such as antibody-affinity chromatography. In antibody-affinity chromatography, a monoclonal antibody is bound to a solid substrate and exposed to a solution containing the desired antigen. The antigen is removed from the solution through an immunospecific reaction with the bound antibody. The polypeptide or polynucleotide is then easily removed from the substrate and purified.

10

15

#### Screening Assays

In yet another aspect, the present invention contemplates a method of screening substances for their ability to affect or modulate the biological activity of *piwi* family gene products. The present invention also contemplates a method of screening substances for their ability to affect or modulate the biological activity of *piwi* family gene products to thereby affect or modulate the growth, proliferation and self-renewing division of stem cells, particularly germline stem cells, and proliferation of primordial germ cells. Utilizing the methods and compositions of the present invention, screening assays for the testing of candidate substances can be derived. A candidate substance is a substance which potentially can promote or inhibit the biological activity of a *piwi* family gene product by binding, or other intermolecular interaction, with the *piwi* family gene product.

20

25

An exemplary method of screening candidate substances for their ability to modulate *piwi* family biological activity comprises the steps of: (a) establishing replicate test and control samples that comprise a biologically

30

active *piwi* family polypeptide; (b) administering a candidate substance to test sample but not the control sample; (c) measuring the biological activity of the *piwi* family polypeptide in the test and the control samples; and (d) determining that the candidate substance modulates *piwi* family biological activity if the biological activity of the *piwi* family polypeptide measured for the test sample is greater or less than the biological activity of the *piwi* family polypeptide measured for the control sample. The biological activities of *piwi* family that may optionally be examined in connection with a screening assay of the present invention comprise modulating the growth, proliferation and self-renewing division of stem cells, particularly germline stem cells, proliferation of primordial germ cells or other biological activity in accordance with the present invention.

The replicate test and control samples can further comprise a cell that expresses a biologically active *piwi* family polypeptide. The present invention also contemplates a recombinant cell line suitable for use in the exemplary method. A candidate substance identified according to the screening assay described herein is contemplated to have the ability to modulate *piwi* family biological activity. Such as candidate compound has utility in the treatment of disorders and conditions associated with the biological activity of *piwi* family.

In a cell-free system, the method comprises the steps of establishing a control system comprising a *piwi* family polypeptide and a ligand wherein the *piwi* family polypeptide is capable of binding to the ligand; establishing a test system comprising the *piwi* family polypeptide, the ligand, and a candidate compound; measuring the binding affinity of the *piwi* family polypeptide and the ligand in the control and the test systems; and determining that the candidate compound modulates *piwi* family activity in a cell-free system if the binding affinity measured for the test system is less than or greater than the binding affinity measured for the control system. An exemplary ligand comprises a monoclonal antibody.

A screening assay of the present invention may also involve determining the ability of a candidate substance to modulate, i.e. inhibit or promote *piwi* family biological activity and preferably, to thereby modulate the growth,

proliferation and self-renewing division of stem cells, particularly germline stem cells, proliferation of primordial germ cells or division of other target cells. Target cells can be either naturally occurring cells known to contain a polypeptide of the present invention or transformed cell produced in accordance with a method of transformation set forth hereinbefore.

As is well known in the art, a screening assay provides a cell under conditions suitable for testing the modulation of *piwi* family biological activity or of growth, proliferation and self-renewing division of stem cells, particularly germline stem cells, or of proliferation of primordial germ cells. These conditions include but are not limited to pH, temperature, tonicity, the presence of relevant metabolic factors (e.g., metal ions such as for example  $\text{Ca}^{++}$ , growth factor, interleukins, or colony stimulating factors), and relevant modifications to the polypeptide such as glycosylation or prenylation. It is contemplated that a polypeptide of the present invention can be expressed and utilized in a prokaryotic or eukaryotic cell. The host cell can also be fractionated into sub-cellular fractions where the receptor can be found. For example, cells expressing the polypeptide can be fractionated into the nuclei, the endoplasmic reticulum, vesicles, or the membrane surfaces of the cell.

In a preferred embodiment, temperature is from about 20°C. to about 50°C., more preferably from about 30°C. to about 40°C. and, even more preferably about 37°C. Osmolality is preferably from about 5 milliosmols per liter (mosm/L) to about 400 mosm/l and, more preferably from about 200 milliosmols per liter to about 400 mosm/l and, even more preferably from about 290 mosm/L to about 310 mosm/L. The presence of factors can be required for the proper testing of *piwi* family biological activity modulation in specific cells. Such factors include, for example, the presence and absence (withdrawal) of growth factor, interleukins, or colony stimulating factors. U.S. Patent Nos. 5,837,479; 5,645,999; 5,786,152; 5,739,278; and 5,352,660 also describe exemplary screening assays, and the entire contents of each are herein incorporated by reference.

In one embodiment, a screening assay is designed to be capable of discriminating candidate substances having selective ability to interact with one

or more of the polypeptides of the present invention but which polypeptides are without a substantially overlapping activity with another of those polypeptides identified herein. Exemplary assays including genetic screening assays and molecular biology screens such as a yeast two-hybrid screen which will effectively identify *piwi*-interacting genes important for *piwi*-mediated cell growth, proliferation and self-renewing division processes. One version of the yeast two-hybrid system has been described (Chien et al., 1991, *Proc. Natl. Acad. Sci. USA*, 88:9578-9582) and is commercially available from Clontech, Palo Alto, California.

#### 10 Therapeutic Methods

As used herein, the terms "*piwi* family activity" and "*piwi* family biological activity" are meant to be synonymous and are meant to refer to any biological activity of a *piwi* family polypeptide. Exemplary biological activities of *piwi* family comprise activity in the modulation of growth, proliferation and self-renewing division of stem cells (including germline and somatic stem cells), cell proliferation (including activity in embryonic development as described in the Examples), proliferation of primordial germ cells or other biological activity in accordance with the present invention. The biological activity can be accomplished by endogenous *piwi* family polypeptides or by *piwi* family polypeptide administered to a subject. Indeed, an isolated and purified *piwi* family polypeptide, recombinant *piwi* family polypeptide, and/or *piwi* family analog or peptidomimetic, each prepared as described herein, can administered to a subject to impart *piwi* family biological activity in the subject and to treat a disorder associated with *piwi* family biological activity in the subject. In such case the imparted *piwi* family biological activity comprises a *piwi* family biological activity in accordance with the therapeutic methods of the present invention.

With respect to the therapeutic methods of the present invention, any subject can be treated, including animal and plant subjects. A preferred subject is a vertebrate subject. Preferred examples of vertebrates include fish and warm-blooded vertebrates. A preferred example of a warm-blooded vertebrate is a mammal. A preferred example of a mammal is a human.



-51-

Additionally, as used herein and in the claims, the term "patient" is contemplated to include both human and animal patients, and thus, veterinary therapeutic uses are contemplated in accordance with the present invention.

Accordingly, the methods of the present invention are particularly  
5 contemplated to be useful in the treatment of warm-blooded vertebrates. Therefore, the invention concerns mammals and birds.

Contemplated is the treatment of mammals such as humans, as well as those mammals of importance due to being endangered (such as Siberian tigers), of economical importance (animals raised on farms for consumption by  
10 humans) and/or social importance (animals kept as pets or in zoos) to humans, for instance, carnivores other than humans (such as cats and dogs), swine (pigs, hogs, and wild boars), ruminants (such as cattle, oxen, sheep, giraffes, deer, goats, bison, and camels), and horses. Also contemplated is the treatment of birds, including the treatment of those kinds of birds that are  
15 endangered, kept in zoos, as well as fowl, and more particularly domesticated fowl, i.e., poultry, such as turkeys, chickens, ducks, geese, guinea fowl, and the like, as they are also of economical importance to humans. Thus, contemplated is the treatment of livestock, including, but not limited to, domesticated swine (pigs and hogs), ruminants, horses, poultry, and the like.

#### 20 Modulation of *piwi* family biological activity

In one embodiment, a therapeutic method according to the present invention comprises administering to a subject a substance that inhibits or promotes *piwi* family biological activity. Such a substance may be identified according to the screening assay set forth above. The method comprises  
25 treating a subject suffering from a disorder associated with or mediated by *piwi* family biological activity by administering to the subject an effective *piwi* family biological activity-modulating amount of a substance identified according to a screening assay described above. By the term "modulating", it is contemplated that the substance can either promote or inhibit the biological activity of *piwi*  
30 family polypeptides, depending on the disorder to be treated. For example, the administering of antibodies against a chosen PIWI family protein (e.g. HIWI,

MIWI, etc.) provides treatment of germline tumor or other developmental or tissue defects caused by *piwi* family-mediated mechanisms.

Insofar as a *piwi* family biological activity modulator can take the form of a polypeptide or of an anti-*piwi* family monoclonal antibody, or fragment thereof, it is to be appreciated that the potency, and therefore an expression of a "therapeutically effective" amount can vary. However, as shown by the present assay methods, one skilled in the art can readily assess the potency of a candidate *piwi* family biological activity modulator of this invention. A *piwi* family biological activity modulator can be measured by a variety of means including through the use of a responsive reporter, which drives expression of a reporter gene; interaction of *piwi* family polypeptides with a monoclonal antibody as described herein; and the like assays.

A preferred *piwi* family biological activity modulator has the ability to substantially interact with *piwi* family in solution at modulator concentrations of less than one (1) micro molar ( $\mu\text{M}$ ), preferably less than 0.1  $\mu\text{M}$ , and more preferably less than 0.01  $\mu\text{M}$ . By "substantially" is meant that at least a 50 percent modulation in *piwi* family biological activity is observed in the presence of the *piwi* family biological activity modulator, and a 50% modulation is referred to herein as an IC50 value.

A therapeutically effective amount of a *piwi* family biological activity modulator of this invention in the form of a monoclonal antibody, or fragment thereof, is typically an amount such that when administered in a physiologically tolerable composition is sufficient to achieve a plasma concentration of from about 0.01 microgram ( $\mu\text{g}$ ) per milliliter (ml) to about 100  $\mu\text{g/ml}$ , preferably from about 1  $\mu\text{g/ml}$  to about 5  $\mu\text{g/ml}$ , and usually about 5  $\mu\text{g/ml}$ .

A therapeutically effective amount of a *piwi* family biological activity modulator of this invention in the form of a polypeptide is typically an amount of polypeptide such that when administered in a physiologically tolerable composition is sufficient to achieve a plasma concentration of from about 0.001 microgram ( $\mu\text{g}$ ) per milliliter (ml) to about 10  $\mu\text{g/ml}$ , preferably from about 0.05  $\mu\text{g/ml}$  to about 1.0  $\mu\text{g/ml}$ .

The monoclonal antibodies or polypeptides of the invention can be administered parenterally by injection or by gradual infusion over time. Although the tissue to be treated can typically be accessed in the body by systemic administration and therefore most often treated by intravenous administration of therapeutic compositions, other tissues and delivery means are contemplated where there is a likelihood that the tissue targeted contains the target molecule. Thus, monoclonal antibodies or polypeptides of the invention can be administered intravenously, intraperitoneally, intramuscularly, subcutaneously, intra-cavity, transdermally, and can be delivered by peristaltic means.

The therapeutic compositions containing a monoclonal antibody or a polypeptide of this invention are conventionally administered intravenously, as by injection of a unit dose, for example. The term "unit dose" when used in reference to a therapeutic composition of the present invention refers to physically discrete units suitable as unitary dosage for the subject, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent; i.e., carrier or vehicle.

The compositions are administered in a manner compatible with the dosage formulation, and in a therapeutically effective amount. The quantity to be administered depends on the subject to be treated, capacity of the subject's system to utilize the active ingredient, and degree of therapeutic effect desired. Precise amounts of active ingredient required to be administered depend on the judgement of the practitioner and are peculiar to each individual. However, suitable dosage ranges for systemic application are disclosed herein and depend on the route of administration. Suitable regimes for administration are also variable, but are typified by an initial administration followed by repeated doses at one or more hour intervals by a subsequent injection or other administration. Alternatively, continuous intravenous infusion sufficient to maintain concentrations in the blood in the ranges specified for *in vivo* therapies are contemplated. Polypeptide modulators can also be modified as

described herein below with respect to biologically active *piwi* family proteins and protein therapy methods using the same.

#### Monoclonal Antibodies

5 The present invention describes, in one embodiment, *piwi* family modulators in the form of monoclonal antibodies which immunoreact with a *piwi* family polypeptide and bind the *piwi* family polypeptide to modulate biological activity as described herein. The invention also describes above cell lines which produce the antibodies, methods for producing the cell lines, and methods for producing the monoclonal antibodies.

10 The term "antibody or antibody molecule" in the various grammatical forms is used herein as a collective noun that refers to a population of immunoglobulin molecules and/or immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antibody combining site or paratope. An "antibody combining site" is that structural portion of an  
15 antibody molecule comprised of heavy and light chain variable and hypervariable regions that specifically binds antigen.

Exemplary antibodies for use in the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules, single chain immunoglobulins or antibodies, those portions of an  
20 immunoglobulin molecule that contain the paratope, including those portions known in the art as Fab, Fab', F(ab')<sub>2</sub> and F(v), and also referred to as antibody fragments. Indeed, it is contemplated to be within the scope of the present invention that a monovalent modulator may optionally be used in the present method. Thus, the terms "modulate", "modulating", and "modulator"  
25 are meant to be construed to encompass such promotion.

The phrase "monoclonal antibody" in its various grammatical forms refers to a population of antibody molecules that contain only one species of antibody combining site capable of immunoreacting with a particular epitope. A monoclonal antibody thus typically displays a single binding affinity for any  
30 epitope with which it immunoreacts. A monoclonal antibody may therefore contain an antibody molecule having a plurality of antibody combining sites, each immunospecific for a different epitope, e.g., a bispecific monoclonal

antibody. Methods of producing a monoclonal antibody, a hybridoma cell, or a hybridoma cell culture are described above.

It is also possible to determine, without undue experimentation, if a monoclonal antibody has the same (i.e., equivalent) specificity (immunoreaction characteristics) as a monoclonal antibody of this invention by ascertaining whether the former prevents the latter from binding to a preselected target molecule. If the monoclonal antibody being tested competes with the monoclonal antibody of the invention, as shown by a decrease in binding by the monoclonal antibody of the invention in standard competition assays for binding to the target molecule when present in the solid phase, then it is likely that the two monoclonal antibodies bind to the same, or a closely related, epitope.

Still another way to determine whether a monoclonal antibody has the specificity of a monoclonal antibody of the invention is to pre-incubate the monoclonal antibody of the invention with the target molecule with which it is normally reactive, and then add the monoclonal antibody being tested to determine if the monoclonal antibody being tested is inhibited in its ability to bind the target molecule. If the monoclonal antibody being tested is inhibited then, in all likelihood, it has the same, or functionally equivalent, epitopic specificity as the monoclonal antibody of the invention.

An additional way to determine whether a monoclonal antibody has the specificity of a monoclonal antibody of the invention is to determine the amino acid residue sequence of the CDR regions of the antibodies in question. "CDRs" (complementarity determining regions) mean the three subregions of the light or heavy chain variable regions which have hypervariable sequences and form loop structures that are primarily responsible for making direct contact with antigen. Antibody molecules having identical, or functionally equivalent, amino acid residue sequences in their CDR regions have the same binding specificity. Methods for sequencing polypeptides are well known in the art.

The immunospecificity of an antibody, its target molecule binding capacity, and the attendant affinity the antibody exhibits for the epitope, are defined by the epitope with which the antibody immunoreacts. The epitope

-56-

specificity is defined at least in part by the amino acid residue sequence of the variable region of the heavy chain of the immunoglobulin that comprises the antibody, and in part by the light chain variable region amino acid residue sequence. Use of the terms "having the binding specificity of" or "having the binding preference of" indicates that equivalent monoclonal antibodies exhibit the same or similar immunoreaction (binding) characteristics and compete for binding to a preselected target molecule.

Humanized monoclonal antibodies offer particular advantages over murine monoclonal antibodies, particularly insofar as they can be used therapeutically in humans. Specifically, human antibodies are not cleared from the circulation as rapidly as "foreign" antigens, and do not activate the immune system in the same manner as foreign antigens and foreign antibodies. Methods of preparing "humanized" antibodies are generally well known in the art, and can readily be applied to the antibodies of the present invention. Thus, the invention contemplates, in one embodiment, a monoclonal antibody of this invention that is humanized by grafting to introduce components of the human immune system without substantially interfering with the ability of the antibody to bind antigen.

The use of a molecular cloning approach to generate antibodies, particularly monoclonal antibodies, and more particularly single chain monoclonal antibodies, is also contemplated. The production of single chain antibodies has been described in the art, see e.g., U.S. Patent No. 5,260,203, the contents of which are herein incorporated by reference. For this, combinatorial immunoglobulin phagemid libraries are prepared from RNA isolated from the spleen of the immunized animal, and phagemids expressing appropriate antibodies are selected by panning on endothelial tissue. The advantages of this approach over conventional hybridoma techniques are that approximately  $10^4$  times as many antibodies can be produced and screened in a single round, and that new specificities are generated by H and L chain combination in a single chain, which further increases the chance of finding appropriate antibodies. Thus, an antibody of the present invention, or a "derivative" of an antibody of the present invention pertains to a single

polypeptide chain binding molecule which has binding specificity and affinity substantially similar to the binding specificity and affinity of the light and heavy chain aggregate variable region of an antibody described herein.

#### Other Modulators

5           Given the disclosure of the *piwi* family activity in tissues herein, it is also contemplated that as yet undefined chemical compounds may be used to modulate *piwi* family activity in tissues in accordance with the methods of the present invention. The identification of such compounds is facilitated by the description of screening assays directed to *piwi* family activity in tissues  
10           presented above.

#### Gene Therapy

          In accordance with the present invention, *piwi* family genes can be used as a tool of gene therapy in humans and mammalian, avian, and fish systems to increase the number of stem cells and to increase proliferation of primordial  
15           germ cells *in vivo*.

          For example, over-proliferation of malignant stem cells is the leading cause of cancer while under-proliferation of stem cells or stem-like progenitor cells leads to tissue dystrophy, anemia, immunodeficiency, and male infertility. The crucial role of stem cells has long been attributed to their ability to self-renew and to generate immense number of specialized cells on demand.  
20           Thus, *piwi* family genes can be used for gene therapy in accordance with the present invention. Exemplary gene therapy methods, including liposomal transfection of nucleic acids into host cells, are described in U.S. Patent Nos. 5,279,833; 5,286,634; 5,399,346; 5,646,008; 5,651,964; 5,641,484; and  
25           5,643,567, the contents of each of which are herein incorporated by reference.

          Briefly, *piwi* family gene therapy directed toward modulation of growth, proliferation and self-renewing division in a target cell is described. Target cells include but are not limited to germline stem cells, primordial germ cells and cancerous or tumorous cells. In one embodiment, a therapeutic method of the  
30           present invention contemplates a method for modulating of growth, proliferation or self-renewing division in a cell comprising the steps of: (a) delivering to the cell an effective amount of a DNA molecule comprising a polynucleotide that

encodes a *piwi* family polypeptide that modulates growth, proliferation or self-renewing division; and (b) maintaining the cell under conditions sufficient for expression of said polypeptide.

In a preferred embodiment, the *piwi* family polypeptide is HIWI or MIWI.  
5 Delivery is preferably accomplished by injecting the DNA molecule into the cell. Where the cell is in a subject delivering is preferably administering the DNA molecule into the circulatory system of the subject. In a preferred embodiment, administering comprises the steps of: (a) providing a vehicle that contains the DNA molecule; and (b) administering the vehicle to the subject.

10 A vehicle is preferably a cell transformed or transfected with the DNA molecule or a transfected cell derived from such a transformed or transfected cell. An exemplary and preferred transformed or transfected cell is a spermatogonial cell or a tumor cell from the tumor being treated. Techniques for transforming or transfecting a cell with a DNA molecule of the present  
15 invention are set forth above.

Alternatively, the vehicle is a virus or an antibody that specifically infects or immunoreacts with an antigen of the tumor. Retroviruses used to deliver the constructs to the host target tissues generally are viruses in which the 3'-LTR (linear transfer region) has been inactivated. That is, these are enhancerless  
20 3'-LTR's, often referred to as SIN (self-inactivating viruses) because after productive infection into the host cell, the 3'-LTR is transferred to the 5'-end and both viral LTR's are inactive with respect to transcriptional activity. A use of these viruses well known to those skilled in the art is to clone genes for which the regulatory elements of the cloned gene are inserted in the space  
25 between the two LTR's. An advantage of a viral infection system is that it allows for a very high level of infection into the appropriate recipient cell.

Antibodies have been used to target and deliver DNA molecules. An N-terminal modified poly-L-lysine (NPLL)-antibody conjugate readily forms a complex with plasmid DNA. A complex of monoclonal antibodies against a cell  
30 surface thrombomodulin conjugated with NPLL was used to target a foreign plasmid DNA to an antigen-expressing mouse lung endothelial cell line and



mouse lung. Those targeted endothelial cells expressed the product encoded by that foreign DNA.

It is also envisioned that this embodiment of the present invention can be practiced using alternative viral or phage vectors, including retroviral  
5 vectors, adenoviral vectors and vaccinia viruses whose genome has been manipulated in alternative ways so as to render the virus non-pathogenic. Methods for creating such a viral mutation are set forth in detail in U.S. Pat. No. 4,769,331, incorporated herein by reference.

By way of specific example, *hiwi*, *miwi*, and their homologs from warm-blooded vertebrates are introduced into isolated spermatogonial cells or other  
10 relevant cells. The re-injection of the transgene-carrying cells into the testis or other relevant tissues provides a treatment for male infertility or other relevant diseases in human and animals.

#### Protein Therapy

15 In another embodiment, the direct introduction of the *piwi* family proteins into a diseased tissue is contemplated to provide a therapeutic effect in human and animals. Such a therapeutic method comprises administering to a subject a therapeutic composition which comprises a biologically active *piwi* family polypeptide of the present invention in amount effective to modulate a *piwi*-  
20 mediated biological activity in the subject.

In one embodiment, a polypeptide for use in such a composition comprises no more than about 100 amino acid residues, preferably no more than about 60 residues, more preferably no more than about 30 residues. Peptides can be linear or cyclic.

25 A subject polypeptide includes any analog, fragment or chemical derivative of a *piwi* family polypeptide. Such a polypeptide can be subject to various changes, substitutions, insertions, and deletions where such changes provide for certain advantages in its use. In this regard, an *piwi* family polypeptide for use in a therapeutic method of the present invention  
30 corresponds to, rather than is identical to, the sequence of a native *piwi* family polypeptide where one or more changes are made and it retains the ability to function as a *piwi* family polypeptide in one or more of the assays as defined

herein. Thus, a polypeptide can be in any of a variety of forms of peptide derivatives, that include amides, conjugates with proteins, cyclized peptides, polymerized peptides, analogs, fragments, chemically modified peptides, and the like derivatives.

5           The term "analog" includes any polypeptide having an amino acid residue sequence substantially identical to a sequence of an endogenous *piwi* family polypeptide in which one or more residues have been conservatively substituted with a functionally similar residue and which displays the *piwi* family biological activity as described herein. Examples of conservative substitutions  
10       include the substitution of one non-polar (hydrophobic) residue such as isoleucine, valine, leucine or methionine for another; the substitution of one polar (hydrophilic) residue for another such as between arginine and lysine, between glutamine and asparagine, between glycine and serine; the substitution of one basic residue such as lysine, arginine or histidine for  
15       another; or the substitution of one acidic residue, such as aspartic acid or glutamic acid for another. Such substitutions are described in detail above with respect to the isolated and purified *piwi* family polypeptide of the present invention.

20           The phrase "conservative substitution" also includes the use of a chemically derivatized residue in place of a non-derivatized residue provided that such polypeptide displays the requisite biological activity.

          "Chemical derivative" refers to a subject polypeptide having one or more residues chemically derivatized by reaction of a functional side group. Such derivatized molecules include for example, those molecules in which free  
25       amino groups have been derivatized to form amine hydrochlorides, p-toluene sulfonyl groups, carbobenzoxy groups, t-butyloxycarbonyl groups, chloroacetyl groups or formyl groups. Free carboxyl groups may be derivatized to form salts, methyl and ethyl esters or other types of esters or hydrazides. Free hydroxyl groups may be derivatized to form O-acyl or O-alkyl derivatives. The  
30       imidazole nitrogen of histidine may be derivatized to form N-im-benzylhistidine. Also included as chemical derivatives are those peptides which contain one or more naturally occurring amino acid derivatives of the twenty standard amino

acids. For examples: 4-hydroxyproline may be substituted for proline; 5-hydroxylysine may be substituted for lysine; 3-methylhistidine may be substituted for histidine; homoserine may be substituted for serine; and ornithine may be substituted for lysine. Polypeptides of the present invention also include any polypeptide having one or more additions and/or deletions or residues relative to the sequence of a polypeptide whose sequence is shown herein, so long as the requisite activity is maintained.

The term "fragment" refers to any subject polypeptide having an amino acid residue sequence shorter than that of a polypeptide whose amino acid residue sequence is shown herein.

When a polypeptide of the present invention has a sequence that is not identical to the sequence of a *piwi* family endogenous polypeptide, it is typically because one or more conservative or non-conservative substitutions have been made, usually no more than about 30 number percent, and preferably no more than 10 number percent of the amino acid residues are substituted. Additional residues may also be added at either terminus of a polypeptide for the purpose of providing a "linker" by which the polypeptides of this invention can be conveniently affixed to a label or solid matrix, or carrier. Labels, solid matrices and carriers that can be used with the polypeptides of this invention are described elsewhere herein.

Amino acid residue linkers are usually at least one residue and can be 40 or more residues, more often 1 to 10 residues, but do not form *piwi* family polypeptide epitopes. Typical amino acid residues used for linking are tyrosine, cysteine, lysine, glutamic and aspartic acid, or the like. In addition, a subject polypeptide can differ, unless otherwise specified, from the natural sequence of a *piwi* family polypeptide by the sequence being modified by terminal-NH<sub>2</sub> acylation, e.g., acetylation, or thioglycolic acid amidation, by terminal-carboxylamidation, e.g., with ammonia, methylamine, and the like terminal modifications. Terminal modifications are useful, as is well known, to reduce susceptibility by proteinase digestion, and therefore serve to prolong half life of the polypeptides in solutions, particularly biological fluids where proteases may be present. In this regard, polypeptide cyclization is also a useful terminal

modification, and is particularly preferred also because of the stable structures formed by cyclization and in view of the biological activities observed for such cyclic peptides as described herein.

Any peptide of the present invention may be used in the form of a pharmaceutically acceptable salt. Suitable acids which are capable of the peptides with the peptides of the present invention include inorganic acids such as trifluoroacetic acid (TFA), hydrochloric acid (HCl), hydrobromic acid, perchloric acid, nitric acid, thiocyanic acid, sulfuric acid, phosphoric acetic acid, propionic acid, glycolic acid, lactic acid, pyruvic acid, oxalic acid, malonic acid, succinic acid, maleic acid, fumaric acid, anthranilic acid, cinnamic acid, naphthalene sulfonic acid, sulfanilic acid or the like. HCl and TFA salts are particularly preferred.

Suitable bases capable of forming salts with the peptides of the present invention include inorganic bases such as sodium hydroxide, ammonium hydroxide, potassium hydroxide and the like; and organic bases such as mono- and tri-alkyl and aryl amines (e.g. triethylamine, diisopropyl amine, methyl amine, dimethyl amine and the like), and optionally substituted ethanolamines (e.g. ethanolamine, diethanolamine and the like).

A peptide of the present invention, also referred to herein as a subject polypeptide, can be synthesized by any of the techniques that are known to those skilled in the polypeptide art, including recombinant DNA techniques. Synthetic chemistry techniques, such as a solid-phase Merrifield-type synthesis, are preferred for reasons of purity, antigenic specificity, freedom from undesired side products, ease of production and the like. An excellent summary of the many techniques available can be found in Steward et al., "Solid Phase Peptide Synthesis", W. H. Freeman Co., San Francisco, 1969; Bodanszky, et al., "Peptide Synthesis", John Wiley & Sons, Second Edition, 1976; J. Meienhofer, "Hormonal Proteins and Peptides", Vol. 2, p. 46, Academic Press (New York), 1983; Merrifield, *Adv Enzymol*, 32:221-96, 1969; Fields et al., *Int. J. Peptide Protein Res.*, 35:161-214, 1990; and U.S. Pat. No. 4,244,946 for solid phase peptide synthesis, and Schroder et al., "The Peptides", Vol. 1, Academic Press (New York), 1965 for classical solution

synthesis, each of which is incorporated herein by reference. Appropriate protective groups usable in such synthesis are described in the above texts and in J. F. W. McOmie, "Protective Groups in Organic Chemistry", Plenum Press, New York, 1973, which is incorporated herein by reference.

5           In general, the solid-phase synthesis methods contemplated comprise the sequential addition of one or more amino acid residues or suitably protected amino acid residues to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid residue is protected by a suitable, selectively removable protecting group. a different, selectively  
10 removable protecting group is utilized for amino acids containing a reactive side group such as lysine.

          Using a solid phase synthesis as exemplary, the protected or derivatized amino acid is attached to an inert solid support through its unprotected carboxyl or amino group. The protecting group of the amino or carboxyl group is then  
15 selectively removed and the next amino acid in the sequence having the complimentary (amino or carboxyl) group suitably protected is admixed and reacted under conditions suitable for forming the amide linkage with the residue already attached to the solid support. The protecting group of the amino or  
20 carboxyl group is then removed from this newly added amino acid residue, and the next amino acid (suitably protected) is then added, and so forth. After all the desired amino acids have been linked in the proper sequence, any remaining terminal and side group protecting groups (and solid support) are removed sequentially or concurrently, to afford the final linear polypeptide.

          The resultant linear polypeptides prepared for example as described  
25 above may be reacted to form their corresponding cyclic peptides. An exemplary method for cyclizing peptides is described by Zimmer et al., *Peptides* 1992, pp. 393-394, ESCOM Science Publishers, B. V., 1993. Typically, tertbutoxycarbonyl protected peptide methyl ester is dissolved in methanol and sodium hydroxide solution are added and the admixture is  
30 reacted at 20°C to hydrolytically remove the methyl ester protecting group. After evaporating the solvent, the tertbutoxycarbonyl protected peptide is extracted with ethyl acetate from acidified aqueous solvent. The

-64-

tertbutoxycarbonyl protecting group is then removed under mildly acidic conditions in dioxane cosolvent. The unprotected linear peptide with free amino and carboxy termini so obtained is converted to its corresponding cyclic peptide by reacting a dilute solution of the linear peptide, in a mixture of  
5 dichloromethane and dimethylformamide, with dicyclohexylcarbodiimide in the presence of 1-hydroxybenzotriazole and N-methylmorpholine. The resultant cyclic peptide is then purified by chromatography.

#### Antisense Oligonucleotide Therapy

In another embodiment, the administering of antisense RNA against a  
10 chosen *piwi* family gene (e.g. *hiwi*, *miwi*, etc.) provides treatment of germline tumor or other developmental or tissue defects caused by *piwi* family-mediated mechanisms. In accordance with the present invention such a therapeutic method may alternatively comprise promoting or inhibiting *piwi* family in a  
15 subject by administering an effective amount of a substance that inhibits or promotes expression of a *piwi* family-encoding nucleic acid segment in the subject. Examples of such a substance, include, for example, an antisense oligonucleotide derived from SEQ ID NOs:1, 3 or 5. Therapeutic methods utilizing antisense oligonucleotides have been described in the art, for example in U.S. Patent Nos. 5,627,158 and 5,734,033, the contents of each of which  
20 are herein incorporated by reference.

#### Formulation of Therapeutic Compositions

The *piwi* family biological activity modulating substances, gene therapy vectors, biologically active *piwi* family gene products, and substances that inhibit or promote expression of a *piwi* family-encoding nucleic acid segment  
25 described above are adapted for administration as a pharmaceutical compositions as described above. Additional formulation and dose preparation techniques have been described in the art, see for example, those described in U.S. Patent No. 5,326,902 issued to Seipp et al. on July 5, 1994, U.S. Patent No. 5,234,933 issued to Mamett et al. on August 10, 1993, and PCT  
30 Publication WO 93/25521 of Johnson et al. published December 23, 1993, the entire contents of each of which are herein incorporated by reference.

For the purposes described above, the identified substances may normally be administered systemically or partially, usually by oral or parenteral administration. The doses to be administered are determined depending upon age, body weight, symptom, the desired therapeutic effect, the route of administration, and the duration of the treatment etc. In a human adult, the doses per person per administration are generally between 1 mg and 500 mg, by oral administration, up to several times per day, and between 1 mg and 100 mg, by parenteral administration up to several times per day. Since the doses to be used depend upon various conditions, as mentioned above, there may be a case in which doses are lower than or greater than the ranges specified above.

Solid compositions for oral administration include compressed tablets, pills, dispersible powders, capsules, and granules. In such compositions, one or more of the active substance(s) is or are, admixed with at least one inert diluent (lactose, mannitol, glucose, hydroxypropylcellulose, microcrystalline cellulose, starch, polyvinylpyrrolidone, magnesium metasilicate alminate, etc.). The compositions may also comprise, as is normal practice, additional substances other than inert diluents: e.g. lubricating agents (magnesium stearate, etc.), disintegrating agents (cellulose, calcium glycolate etc.), and assisting agent for dissolving (glutamic acid, aspartic acid, etc.) stabilizing agent (lactose etc.). The tablets or pills may, if desired, be coated with gastric or enteric material (sugar, gelatin, hydroxypropylcellulose or hydroxypropylmethyl cellulose phthalate, etc.). Capsules include soft ones and hard ones.

Liquid compositions for oral administration include pharmaceutically-acceptable emulsions, solutions, suspensions, syrups and elixirs. In such compositions, one or more of the active substance(s) is or are admixed with inert diluent(s) commonly used in the art (purified water, ethanol etc.). Besides inert diluents, such compositions may also comprise adjuvants (wetting agents, suspending agents, etc.), sweetening agents, flavoring agents, perfuming agents and preserving agents.

Other compositions for oral administration include spray compositions which may be prepared by known methods and which comprise one or more of the active substance(s). Spray compositions may comprise additional substances other than inert diluents: e.g. preserving agents (sodium sulfite, etc.), isotonic buffer (sodium chloride, sodium citrate, citric acid, etc.). For preparation of such spray compositions, for example, the method described in U.S. Patent Nos. 2,868,691 or 3,095,355 may be used.

Injections for parenteral administration include sterile aqueous or non-aqueous solution, suspensions and emulsions. In such compositions, one or more of active substance(s) is or are admixed with at least one inert aqueous diluent(s) (distilled water for injection, physiological salt solution etc.) or inert non-aqueous diluent(s) (propylene glycol, polyethylene glycol, olive oil, ethanol, POLYSOLBATE 80® etc.). Injections may comprise additional other than inert diluents: e.g. preserving agents, wetting agents, emulsifying agents, dispersing agents, stabilizing agents (lactose, etc.), assisting agents such as for dissolving (glutamic acid, aspartic acid, etc.). They may be sterilized, for example, by filtration through a bacteria-retaining filter, by incorporation of sterilizing agents in the compositions or by irradiation. They also be manufactured in the form of sterile solid compositions, for example, by freeze-drying, and which can be dissolved in sterile water or some other sterile diluents for injection immediately before use.

Other compositions for administration include liquids for external use, and endermic linaments (ointment, etc.), suppositories and pessaries which comprise one or more of the active substance(s) and may be prepared by known methods.

#### Detecting a Polynucleotide or a Polypeptide of the Present Invention

Alternatively, the present invention provides a method of detecting a polypeptide of the present invention, wherein the method comprises immunoreacting the polypeptides with antibodies prepared according to the method described above to form antibody-polypeptide conjugates, and detecting the conjugates.



In yet another embodiment, the present invention contemplates a method of detecting messenger RNA transcripts that encode a polypeptide of the present invention, wherein the method comprises hybridizing the messenger RNA transcripts with polynucleotide sequences that encode the polypeptide to form duplexes; and detecting the duplex. Alternatively, the present invention provides a method of detecting DNA molecules that encode a polypeptide of the present invention, wherein the method comprises hybridizing DNA molecules with a polynucleotide that encodes that polypeptide to form duplexes; and detecting the duplexes.

The detection and screening assays disclosed herein can be used as a prognosis tool. *hiwi*, *miwi* and their mammalian homologous genes as well as their protein products can be readily used in clinical setting as a prognostic indicator for screening infertility, germline cancer, or other heritable *piwi*-related diseases in human and animals.

The detection and screening assays disclosed herein can be also used as a part of a diagnostic method. *hiwi*, *miwi*, and their mamalian homologs as well as their proteins can be readily used in clinical practice to diagnose the cause of infertility, testicular cancer, or other heritable *piwi*-related diseases in human and animals.

#### Screening Assays for a Polypeptide of the Present Invention

The present invention provides a method of screening a biological sample for the presence of a *piwi* family polypeptide. Preferably, the *piwi* family polypeptide possesses a biological activity in accordance with the present invention. A biological sample to be screened can be a biological fluid such as extracellular or intracellular fluid or a cell or tissue extract or homogenate. A biological sample can also be an isolated cell (e.g., in culture) or a collection of cells such as in a tissue sample or histology sample. A tissue sample can be suspended in a liquid medium or fixed onto a solid support such as a microscope slide.

In accordance with a screening assay method, a biological sample is exposed to an antibody immunoreactive with the polypeptide whose presence is being assayed. Typically, exposure is accomplished by forming an admixture in a liquid medium that contains both the antibody and the candidate polypeptide. Either the antibody or the sample with the polypeptide can be affixed to a solid support (e.g., a column or a microtiter plate).

The biological sample is exposed to the antibody under biological reaction conditions and for a period of time sufficient for antibody-polypeptide conjugate formation. Biological reaction conditions include ionic composition and concentration, temperature, pH and the like.

Ionic composition and concentration can range from that of distilled water to a 2 molal solution of NaCl. Preferably, osmolality is from about 100 mosmols/l to about 400 mosmols/l and, more preferably from about 200 mosmols/l to about 300 mosmols/l. Temperature preferably is from about 4°C. to about 100°C., more preferably from about 15°C. to about 50°C. and, even more preferably from about 25°C to about 40°C. pH is preferably from about a value of 4.0 to a value of about 9.0, more preferably from about a value of 6.5 to a value of about 8.5 and, even more preferably from about a value of 7.0 to a value of about 7.5. The only limit on biological reaction conditions is that the conditions selected allow for antibody-polypeptide conjugate formation and that the conditions do not adversely affect either the antibody or the polypeptide.

Exposure time will vary *inter alia* with the biological conditions used, the concentration of antibody and polypeptide and the nature of the sample (e.g., fluid or tissue sample). Means for determining exposure time are well known to one of ordinary skill in the art. Typically, where the sample is fluid and the concentration of polypeptide in that sample is about  $10^{-10}$ M, exposure time is from about 10 minutes to about 200 minutes.

The presence of polypeptide in the sample is detected by detecting the formation and presence of antibody-polypeptide conjugates. Means for detecting such antibody-antigen (e.g., receptor polypeptide) conjugates or complexes are well known in the art and include such procedures as

centrifugation, affinity chromatography and the like, binding of a secondary antibody to the antibody-candidate receptor complex.

In one embodiment, detection is accomplished by detecting an indicator affixed to the antibody. Exemplary and well known such indicators include  
5 radioactive labels (e.g.,  $^{32}\text{P}$ ,  $^{125}\text{I}$ ,  $^{14}\text{C}$ ), a second antibody or an enzyme such as horse radish peroxidase. Means for affixing indicators to antibodies are well known in the art. Commercial kits are available.

#### Screening Assay for Anti-Polypeptide Antibody

In another aspect, the present invention provides a method of screening  
10 a biological sample for the presence of antibodies immunoreactive with a *piwi* family polypeptide. Preferably the *piwi* family polypeptide has a biological activity in accordance with the present invention. In accordance with such a method, a biological sample is exposed to a *piwi* family polypeptide under biological conditions and for a period of time sufficient for antibody-polypeptide  
15 conjugate formation and the formed conjugates are detected.

#### Screening Assay for Polynucleotide That Encodes a *piwi* family Polypeptide of the Present Invention

A DNA molecule and, particularly a probe molecule, can be used for hybridizing as an oligonucleotide probe to a DNA source suspected of  
20 encoding a *piwi* family polypeptide of the present invention. Preferably the *piwi* family polypeptide possesses a biological activity in accordance with the present invention. The probing is usually accomplished by hybridizing the oligonucleotide to a DNA source suspected of possessing a *piwi* family gene. In some cases, the probes constitute only a single probe, and in others, the  
25 probes constitute a collection of probes based on a certain amino acid sequence or sequences of the polypeptide and account in their diversity for the redundancy inherent in the genetic code.

A suitable source of DNA for probing in this manner is capable of expressing a polypeptide of the present invention and can be a genomic library  
30 of a cell line of interest. Alternatively, a source of DNA can include total DNA

from the cell line of interest. Once the hybridization method of the invention has identified a candidate DNA segment, one confirms that a positive clone has been obtained by further hybridization, restriction enzyme mapping, sequencing and/or expression and testing.

5           Alternatively, such DNA molecules can be used in a number of techniques including their use as: (1) diagnostic tools to detect normal and abnormal DNA sequences in DNA derived from patient's cells; (2) reagents for detecting and isolating other members of the polypeptide family and related polypeptides from a DNA library potentially containing such sequences; (3)  
10          primers for hybridizing to related sequences for the purpose of amplifying those sequences; (4) primers for altering native *piwi* family DNA sequences; as well as other techniques which rely on the similarity of the DNA sequences to those of the DNA segments herein disclosed.

          As set forth above, in certain aspects, DNA sequence information  
15          provided by the invention allows for the preparation of relatively short DNA (or RNA) sequences (e.g., probes) that specifically hybridize to encoding sequences of a selected *piwi* family gene. In these aspects, nucleic acid probes of an appropriate length are prepared based on a consideration of the encoding sequence for a polypeptide of this invention. The ability of such  
20          nucleic acid probes to specifically hybridize to other encoding sequences lend them particular utility in a variety of embodiments. Most importantly, the probes can be used in a variety of assays for detecting the presence of complementary sequences in a given sample. However, other uses are envisioned, including the use of the sequence information for the preparation of mutant species  
25          primers, or primers for use in preparing other genetic constructions.

          To provide certain of the advantages in accordance with the invention, a preferred nucleic acid sequence employed for hybridization studies or assays includes probe sequences that are complementary to at least a 14 to 40 or so long nucleotide stretch of a nucleic acid sequence of the present invention,  
30          such as a sequence shown in any of SEQ ID NOs:1, 3 and 5. A size of at least 14 nucleotides in length helps to ensure that the fragment is of sufficient length to form a duplex molecule that is both stable and selective. Molecules having

-71-

complementary sequences over stretches greater than 14 bases in length are generally preferred, though, to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having  
5 gene-complementary stretches of 14 to 20 nucleotides, or even longer where desired. Such fragments can be readily prepared by, for example, directly synthesizing the fragment by chemical means, by application of nucleic acid reproduction technology, such as the PCR technology of U.S. Pat. No. 4,683,202, herein incorporated by reference, or by introducing selected  
10 sequences into recombinant vectors for recombinant production.

Accordingly, a nucleotide sequence of the present invention can be used for its ability to selectively form duplex molecules with complementary stretches of the gene. Depending on the application envisioned, one employs varying conditions of hybridization to achieve varying degrees of selectivity of the probe  
15 toward the target sequence. For applications requiring a high degree of selectivity, one typically employs relatively stringent conditions to form the hybrids. For example, one selects relatively low salt and/or high temperature conditions, such as provided by 0.02M-0.15M NaCl at temperatures of 50°C to 70°C. Such conditions are particularly selective, and tolerate little, if any,  
20 mismatch between the probe and the template or target strand.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template or where one seeks to isolate polypeptide coding sequences from related species, functional equivalents, or the like, less stringent hybridization  
25 conditions are typically needed to allow formation of the heteroduplex. Under such circumstances, one employs conditions such as 0.15M-0.9M salt, at temperatures ranging from 20°C to 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions  
30 can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily

-72-

manipulated, and thus will generally be a method of choice depending on the desired results.

In certain embodiments, it is advantageous to employ a nucleic acid sequence of the present invention in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of giving a detectable signal. In preferred embodiments, one likely employs an enzyme tag such a urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmentally undesirable reagents. In the case of enzyme tags, calorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with complementary nucleic acid-containing samples.

In general, it is envisioned that the hybridization probes described herein are useful both as reagents in solution hybridization as well as in embodiments employing a solid phase. In embodiments involving a solid phase, the sample containing test DNA (or RNA) is adsorbed or otherwise affixed to a selected matrix or surface. This fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes under desired conditions. The selected conditions depend *inter alia* on the particular circumstances based on the particular criteria required (depending, for example, on the G+ C contents, type of target nucleic acid, source of nucleic acid, size of hybridization probe, etc.). Following washing of the hybridized surface so as to remove nonspecifically bound probe molecules, specific hybridization is detected, or even quantified, by means of the label.

#### Assay Kits

In another aspect, the present invention contemplates assay kits for detecting the presence of a polypeptide of the present invention in biological samples, where the kits comprise a first container containing a first antibody capable of immunoreacting with the polypeptide, with the first antibody present in an amount sufficient to perform at least one assay. Preferably, the assay kits

-73-

of the invention further comprise a second container containing a second antibody that immunoreacts with the first antibody. More preferably, the antibodies used in the assay kits of the present invention are monoclonal antibodies. Even more preferably, the first antibody is affixed to a solid support. More preferably still, the first and second antibodies comprise an indicator, and, preferably, the indicator is a radioactive label or an enzyme.

The present invention also contemplates a kit for screening agents. Such a kit can contain a polypeptide of the present invention. The kit can contain reagents for detecting an interaction between an agent and a receptor of the present invention. The provided reagent can be radiolabeled. The kit can contain a known radiolabelled agent capable of binding or interacting with a receptor of the present invention.

In an alternative aspect, the present invention provides assay kits for detecting the presence, in biological samples, of a polynucleotide that encodes a polypeptide of the present invention, the kits comprising a first container that contains a second polynucleotide identical or complementary to a segment of at least 10 contiguous nucleotide bases of, as a preferred example, SEQ ID NOs:1, 3 and 5.

In another embodiment, the present invention contemplates assay kits for detecting the presence, in a biological sample, of antibodies immunoreactive with a polypeptide of the present invention, the kits comprising a first container containing a *piwi* family polypeptide, that immunoreacts with the antibodies, with the polypeptide present in an amount sufficient to perform at least one assay. Preferably, the *piwi* family polypeptide has a biological activity in accordance with the present invention. The reagents of the kit can be provided as a liquid solution, attached to a solid support or as a dried powder. Preferably, when the reagent is provided in a liquid solution, the liquid solution is an aqueous solution. Preferably, when the reagent provided is attached to a solid support, the solid support can be chromatograph media or a microscope slide. When the reagent provided is a dry powder, the powder can be reconstituted by the addition of a suitable solvent. The solvent can be provided.

### EXAMPLES

The following examples have been included to illustrate preferred modes of the invention. Certain aspects of the following examples are described in terms of techniques and procedures found or contemplated by the present  
5 inventors to work well in the practice of the invention. These examples are exemplified through the use of standard laboratory practices of the inventors. In light of the present disclosure and the general level of skill in the art, those of skill will appreciate that the following examples are intended to be exemplary only and that numerous changes, modifications and alterations can be  
10 employed without departing from the spirit and scope of the invention.

In exploring stem cell renewal mechanisms in the germline, the *piwi* gene is disclosed herein to be essential for GSC maintenance in *Drosophila*. The role of the *piwi* gene is characterized herein by recessive mutations that cause failure in GSC maintenance in both females and males. *piwi* mutant  
15 gonads contain a normal number of GSCs at the onset of gametogenesis. However, mutant adult gonads are devoid of GSCs and contain only a very small number of gametes that is approximately equal to or less than the wild-type number of GSCs. These observations reveal that *piwi* is essential for GSC maintenance in both males and females.

#### Summary of Observations in Examples

  
20

The purification and isolation of cDNAs representing *piwi* family genes from human, mouse, *Drosophila* are disclosed in the Examples, along with the primary structure of the PIWI proteins in human, mouse and *Drosophila* as deduced from the cDNA sequences. Additionally, the essential function of *piwi*  
25 in *Drosophila* germline stem cell growth, proliferation and self-renewal and in embryonic development is also disclosed.

Further, the identification of *piwi* homologous genes, *prg-1* and *prg-2* in *C. elegans* and the demonstration of the important function of *prg-1* and *prg-2* in *C. elegans* germline self-renewal similar to that of *piwi* in *Drosophila* is described in the Examples. The identification of two *piwi* family-like genes in  
30 *Arabidopsis*, *Zwille* and *argonaute(ago)*, that are known to be required for meristem cell divisions is also described.



The growth, proliferation and self-renewing division modulating biological activity of human and murine *piwi* genes (named *hiwi* and *miwi*, respectively) in germline stem cell renewal, proliferation of primordial germ cells and in other developmental processes is characterized in the Examples based on the following data and reasoning: (i) *hiwi* and *miwi* are more homologous to *piwi* than *prg-I*, *prg-II*, *Zwille* and *ago* to *piwi*; (ii) human and murine germline stem cells are much more similar to *Drosophila* germline cells than to the *C. elegans* germline or *Arabidopsis* meristem cells; (iii) the function of *piwi* family genes are well conserved among *Drosophila*, *C. elegans*, and *Arabidopsis*, which are evolutionarily more diverse than between *Drosophila* and human/mammals; and (iv) *hiwi* and *miwi* is indeed specifically expressed in the adult human and murine testis, respectively, but not in the adult human or murine ovary which does not contain germline stem cells. *miwi* is further shown to be expressed specifically in the germline stem cell region in the mouse testis. Therefore, *hiwi*, *miwi*, and other mammalian *piwi* family genes are shown to play an important role in germline stem cell division and other stem cell-related processes.

The specific expression pattern of *hiwi* and *miwi* in the adult human and murine testis, respectively, but not in the adult human or murine ovary which do not contain germline stem cells is also demonstrated in the Examples. *miwi* is further shown to be expressed in the germline stem cell region in the mouse testis. The discovery of a highly conserved protein signature motif of 43-amino acid residues, the PIWI box, that exists in the *piwi* family genes as well as in a broader class of novel genes which may be structurally and functionally defined by the PIWI box is also described in the Examples. The demonstration that the PIWI protein and, thus, the PRG-1, PRG-2, MIWI, HIWI, Zwille, and AGO proteins are nuclear proteins is also disclosed.

#### Example 1

##### Isolation and Purification of *piwi*

The cloning and characterization of the *Drosophila piwi* gene is disclosed in this Example. Germline stem cells (GSCs) serve as the source for gametogenesis in diverse organisms. The *piwi* gene product is required for the

asymmetric division of GSC to produce and maintain a daughter GSC but is not essential for the further differentiation of the committed daughter cell. Genetic mosaic and RNA *in situ* analyses indicate that *piwi* expression in adjacent somatic cells regulates GSC division.

5           *piwi* encodes a highly basic novel protein well conserved during evolution. This Example also describes the isolation of *piwi* family homologs in *C. elegans* and human. This Example further describes the identification of *Arabidopsis piwi* family-like genes known to be required for meristem cell maintenance. Decreasing *C. elegans piwi* expression reduces the proliferation  
10 of GSC-equivalent cells. Thus, *piwi* represents a novel class of genes required for GSC division in diverse organisms.

*Drosophila* has been used in the art as a model for studying mechanisms that are conserved among diverse developmental systems. As show herein, this is also the case for the study of stem cells. In *Drosophila*,  
15 stem cells exist in the germline at the apical tip of the ovariole, the functional unit of the ovary. In the ovariole, germline stem cells (GSCs) are located in a specialized structure called the germarium, as first proposed by Brown and King (1962, 1964). The existence of GSCs was strongly supported by germline clonal analysis (Schüpbach et al., 1978; Wieschaus and Szabad, 1979) and  
20 directly verified by laser ablation (Lin and Spradling, 1993). In each germarium, 2-3 GSCs contact the somatic basal terminal filament cells. GSCs undergo oriented asymmetric divisions to produce a daughter stem cell which remains associated with the terminal filament and a differentiated daughter, the cystoblast, that becomes displaced one cell away from the terminal filament  
25 (Lin and Spradling, 1997; Deng and Lin, 1997). The germline cyst then becomes enveloped by follicle cells produced by somatic stem cells (Margolis and Spradling, 1995) to form an egg chamber, which buds off the germarium, joins pre-existing egg chambers in a linear array to form an ovariole, and eventually develops into a mature egg (reviewed in King, 1970; Spradling,  
30 1993). This assembly-line organization, with each egg chamber representing a differentiated stem cell product whose position along the ovariole

corresponds to its birth order, provides a unique opportunity to study stem cell division.

To investigate the nature of the *piwi*-mediated mechanism, this Example reports that *piwi* mediates a somatic signaling mechanism essential for the division and maintenance of GSCs in *Drosophila*. Moreover, *piwi* encodes a novel basic protein well conserved in *C. elegans* and humans, and also conserved in *Arabidopsis*, where two *piwi* family-like genes are known to be required for meristem cell maintenance. It is also demonstrated that the *C. elegans piwi* family homolog is also required for germline proliferation and maintenance. Thus, *piwi* family genes represent essential stem cell genes existing in diverse organisms.

*Drosophila* strains and culture. All strains were grown at 25°C on yeast-containing molasses/agar medium. The *piwi*<sup>1</sup>, *piwi*<sup>3</sup> and *piwi*<sup>4</sup> mutations are single PZ insertional mutations reported by Lin and Spradling (1997) while the *piwi*<sup>2</sup> mutation is a single P-ry<sup>11</sup> transposable insertion. *Oregon R* (*Ore-R*) serves as the wildtype strain for all experiments.

Phenotypic analysis of *piwi* mutants. *Piwi* mutations were balanced over *CyO-y*<sup>+</sup> and introduced into the *y/y* genetic background. The homozygous and heterozygous larvae were separated at the late third instar stage by scoring mouth hooks and denticle belts for *y*. They were then dissected to obtain the ovary immediately or were aged at 25°C to desired pupal stages before dissection. Dissected ovaries were analyzed by immunofluorescence microscopy for defects in ovary differentiation and GSC division.

Immunocytochemistry and immunofluorescence microscopy. Wild type and mutant ovaries from larval, pupal, or adult stages were dissected, fixed and stained as described by Lin et al. (1994). For immunofluorescence staining, anti-VASA antibodies (Hay et al., 1990) were used to specifically mark germ cells at 1:2000 dilution. Anti- $\alpha$ -spectrin antibodies (Byers et al., 1987) were used to outline somatic cells and to mark spectrosomes and fusomes (Lin and Spradling, 1995) at 1:200 dilution. The monoclonal anti-MYC epitope antibody 1-9E10.2 was described by Evan et al. (1985) and was used at 1:2 dilution. All the fluorescence-conjugated secondary antibodies were from Jackson

-78-

ImmunoResearch Laboratory Inc., Westgrove, Pennsylvania, and were used at 1:200 dilution.

Immunofluorescently labeled samples were also counter-stained with the DNA specific dye DAPI as described in Lin and Spradling (1993). The immunologically labeled samples were examined using Nomarski and epifluorescence microscopy under a Zeiss AXIOPLAN® microscope (Zeiss, Oberkochen, Germany) equipped with a STAR-1™ cooled CCD camera (Photometrics, Tucson, Arizona). Selected samples were further analyzed by confocal microscopy using a Zeiss LSM410™ confocal microscope mounted on an AXIOVERT® 100. Images from the Zeiss AXIOPLAN® were collected using IP Lab software and confocal images were processed using the LSM410 software. All images were processed using the Adobe Photoshop™ program (Adobe Systems, Inc., San Jose, California).

Cloning of *piwi* and mapping of its mutations. *piwi* maps to the left arm of the second chromosome within polytene band 32C, as defined by insertions of single transposable elements (Lin and Spradling, 1997). DNA fragments flanking the *piwi*<sup>1</sup>, *piwi*<sup>3</sup>, and *piwi*<sup>4</sup> insertions were recovered by a plasmid rescue method (Wilson *et al.*, 1989). Genomic DNA in the rescued plasmid was used to isolate additional genomic clones from a λ genomic DNA library (Maniatis *et al.*, 1978) to constitute a 36 kb region of overlapping walk clones (Figure 2). The insertion sites for *piwi*<sup>1</sup>, *piwi*<sup>3</sup>, and *piwi*<sup>4</sup> were first mapped by Southern blots of EcoRI digested walk clones, and then precisely determined by DNA sequencing. The site of the *piwi*<sup>2</sup> insertion was determined by sequencing its flanking DNA fragments isolated by inverse PCR of *piwi*<sup>2</sup> genomic DNA using primers specific to the P-element termini (Ochman *et al.*, 1990).

Genomic DNA fragments derived from the *piwi* locus were used to probe Northern blots of poly (A)+ RNA from wild-type male and female flies. Three non-overlapping transcripts, α (1.1 kb), β (0.64 kb), and c12 (3.1 kb) were identified in the region and only c12 was interrupted by *piwi* mutations. 9.36 kb of genomic DNA that encompasses the c12 transcription unit was sequenced either using a Sequenase kit (U.S. Biochemical) or an ABI 377

Prism DNA Sequencer (Perkin-Elmer, Applied Biosystems Inc). All sequence data was analyzed using the DNASTAR software package (DNASTAR, Madison, WI). Two additional transcripts, LD08487 (1.43 kb) and LD13681, were identified on the right side of the P-insertions by BLAST analysis (Altschul, et al., 1990) of genomic sequence.

Two overlapping cDNA clones which together represent the 3.1 kb full length c12 mRNA and the LD08487 cDNA were sequenced. Comparison between the genomic and the cDNA sequences precisely determined the intron-exon organization of these two transcription units. The LD13681 transcription unit were also mapped by partial sequencing. Standard molecular biology techniques were carried out according to Sambrook et al. (1989).

Transgenic rescue. The 6.8 kb *HindIII/EcoRV* (pRc12; *piwi*) and 3.9 kb *PstI* (pRLD; LD08487) genomic fragments were subcloned into the P element transformation vector pCaSpeR4 (Pirrotta, 1988). Transgenic flies were produced according to Spradling and Rubin (1982), using *w; Δ2-3 Sb e/TM6 e* (Robertson et al., 1988) embryos as recipients. Transgenes were then separated from the Δ2-3 transposase and introduced into the homozygous *piwi*<sup>1</sup> background for fertility tests by genetic crosses. Six independent pRc12 inserts and eight independent pRLD inserts were recovered. Rescue crosses were carried out at 25°C.

Northern blot analysis and *piwi* cDNA isolation. Poly (A)<sup>+</sup> mRNA isolation, Northern blot preparation, and <sup>32</sup>P-labeled probe preparation by random-hexamer extension were according to Sambrook et al. (1989). For Fig. 2B, the probe was made from a 5.4 kb *HindIII* genomic fragment encompassing the *piwi* transcription unit. The same probe was used to isolate a 2.1 kb partial cDNA from the pNB40 third instar library of Brown and Kafatos, 1989. Using the sites within the polylinker (*HindIII* at 5' end and *NotI* at 3' end), the 2.1 kb cDNA was subcloned into pBlueScript KS(+) (Stratagene), resulting in a construct denoted pDC1. A primer, pDC1 5'REV, 5'-ACGATAAGTTCTGTTAT-3' (SEQ ID NO:11) was designed downstream of the 5' end of pDC1 and was used in combination with library specific primers to screen the Tulle Hazelrigg ovarian cDNA library (Friedman et al., 1990). A

partially overlapping 500 bp PCR product was isolated, cloned into the pGEM-T vector (Promega, Madison, Wisconsin), and used to probe the same library. An overlapping 1 kb cDNA clone was isolated and subcloned into the *Sall* site of pDC1 to produce a full length 3.1 kb *piwi* cDNA construct denoted pDC2.

5        Characterization of *piwi* homologs from *C. elegans* and *Homo sapiens*.

The BLAST algorithm (Altschul et al., 1990) was used to search the non-redundant databases at the National Center for Biotechnology Information (NCBI) at the National Library of Medicine. The Prosite Protein Motif Database (DNASstar, Madison, WI) was also searched to identify domains shared  
10        between *piwi* and other known proteins. The PIWI ORF was used to query the SWISS-Prot database by BLAST for homologous proteins. Significantly, the search identified two *C. elegans* predicted proteins of unknown function. Expressed sequence tags (ESTs) corresponding to these two gene products were then identified as D2030.6 (*prg-1*; GenBank accession No. Z73906) and  
15        C01G5.2 (*prg-2*; GenBank accession NoU50068) from GenBank. Dr. Yuji Kohara (National Institute of Genetics, Japan) kindly provided  $\lambda$  ZAPII phagemid clones corresponding to *prg-1* and *prg-2*.

The *piwi* cDNA was further used to query the NCBI dbEST database and identified an EST, zw68h01.r1, isolated from a human testes cDNA library,  
20        GenBank accession number AA430311. This 0.9 kb EST clone was sequenced and used to screen an oligo-(dT) primed human testis cDNA library (Clontech, Palo Alto, California). A 2.3 kb cDNA clone corresponding to *hiwi* was isolated and sequenced.

Whole-mount tissue *in situ* hybridization. Ovarian whole-mount tissue  
25        *in situ* hybridization was performed as described in Yue and Spradling (1992) with DIG-labeled (Genius Kit, Boehringer Mannheim) full length *piwi* cDNA from pDC2 as a probe. For detecting RNA in surface cells, ovaries were digested 5 minutes with 50  $\mu$ g/ml proteinase K. Under this condition, only terminal filament cells and anterior sheath cells, but not any other surface cells, were  
30        stained with the *piwi* cDNA probe. The pBST control probe did not show any significant signal. For detecting RNA in inner germarial cells, the digestion was 7-10 minutes. Under this condition, the *piwi* probe revealed a hybridization

pattern as shown in Results, the pBST control probe did not show any significant signal, while the positive control *oskar* probe showed specific staining in presumptive oocytes (Ephrussi et al., 1991). Staging of egg chambers is according to King (1970).

5        Genetic clonal analyses and germarial transplantation. The *piwi*<sup>1</sup> and *piwi*<sup>2</sup> germline clones were generated with the FLP-DFS technique as described in Chou and Perrimon (1996). *y, w P[hsFLP]12; P[ovoD1]2L, FRT40A/CyO* males were crossed to *w; piwi*<sup>1</sup>, *FRT40A/CyO* virgin females to produce *y,w P[hsFLP]12; P[ovoD1]2L, FRT40A/piwi*<sup>1</sup>, *FRT40A* progeny.  
10       Identical crosses were performed with *w; piwi*<sup>2</sup>, *FRT40A/CyO* virgin females. The genotype of the flies were verified both by Southern analysis for the presence of *piwi* P insertion and the FRT as well as by sterility tests by backcrossing to the original *piwi* allele. Crosses were carried out for two days to produce transheterozygous progeny. After 2 days, adults were transferred  
15       to fresh vials. Larvae were heat shocked on days 3 and 4 for one hour in a 37°C water bath to induce mitotic recombination. The heat-shocked females with germline clones were crossed to *Ore-R* males for fertility tests and their ovarian phenotype examined by immunofluorescence microscopy.

20       To generate somatic clones of *piwi*<sup>1</sup> and *piwi*<sup>2</sup> the above described *piwi*<sup>1</sup>, *FRT40A/CyO* and *piwi*<sup>2</sup>, *FRT40A/CyO* strains were crossed to *w; P[hsFLP]<sup>9F</sup>; P[w<sup>+</sup>; hs-NM], FRT40A/CyO* males as described in Xu and Rubin (1993). Females of the genotype *w, P[hsFLP]<sup>9F</sup>; P[w<sup>+</sup>; hs-NM], FRT40A/piwi, FRT40A* were crossed to *Ore-R* males for fertility tests. Clonal induction was carried out as described above except that females of the appropriate genotype  
25       were subjected to an additional 20-minute heat shock at 37°C 90 minutes prior to dissection to induce the expression of the myc tag. Germarial transplantation was carried out as described by Lin and Spradling (1993).

30       RNA-mediated interference (RNAi). Templates for *in vitro* transcription were cloned into pBluescript KS (+) and transcribed using either T7 or T3 RNA polymerases according to the manufacturer's protocol (Megascript T7 and T3 kits, Ambion). *prg-1* cDNA served as the experimental RNA for injection while pBluescript KS(+) RNA was used as a negative control. RNA integrity was

-82-

determined by gel electrophoresis; concentration was determined by a combination of UV spectrophotometry and ethidium bromide staining. Injections used an equimolar mixture of uncapped sense and antisense RNA at a concentration of 10 mg/ml. Gonadal injections of wild type N2 Bristol hermaphrodites was done as described in Mello and Fire (1995).

After recovery and transfer to standard solid media, injected animals were transferred to fresh culture plates at 24 hours intervals to facilitate the identification of phenotypic differences. Germlines of progeny produced between 6 and 48 hours after injection were examined for gamete differentiation either by DAPI staining or fertility. For DAPI analysis, adult hermaphrodites displaying RNAi induced phenotypes were fixed overnight in Carnoy's solution (60% ethanol; 30% acetic acid; 10% chloroform; Sulston and Hodgkin, 1988). The mitotic index is defined as the ratio between the number of prophase, metaphase, anaphase, and telophase nuclei and the total number of nuclei in the MPZ.

Piwi mutations eliminate the growth, proliferation and self-renewing division of germline stem cells. *piwi*<sup>1</sup> mutant ovaries contain a normal number of mispositioned GSCs at the onset of oogenesis at the late third instar larval stage, which, however, leads to an equal or somewhat smaller number of gametes in the adult gonads that no longer contain GSCs (Lin and Spradling, 1997). This failure of germline maintenance could be due to: (1) the differentiation of GSCs without growth, proliferation and self-renewing divisions; (2) a defect in the asymmetry of GSC division, producing aberrant germ cells that eventually degenerate; and/or (3) a secondary defect influenced by abnormal ovary differentiation.

To examine whether the failure of germline maintenance is indeed a secondary defect due to abnormal ovary development, the ovarian morphology of *piwi*<sup>2</sup> and *piwi*<sup>3</sup> mutants which also fail to maintain GSCs was examined. The examination relied on Nomarski optics as well as markers that specifically identify germ cells, somatic cells, and spectrosomes/fusomes, germline specific organelles that indicate individual stages of germ cell development (Lin et al., 1994; Lin and Spradling, 1995). These mutant ovaries show normal



morphology at the third instar larval stage. Their germline cells are normal in number and are correctly positioned along the medial plane of the ovary. Moreover, the expected number of terminal filaments are forming, so that at the pupal stage, the ovary differentiates normally, partitioning GSCs and their daughter cells correctly into individual germaria and ovarioles. GSCs are able to divide several times to provide a normal complement of germ cells to the germarium. Yet, GSCs subsequently fail to continue self-renewing divisions, and the existing germline cysts often degenerate during the late pupal stage so that the adult ovarioles contain germlineless germaria and fewer egg chambers than expected. These observations indicate that the failure of GSC maintenance in *piwi* mutants is not a secondary defect due to abnormal ovary development.

To examine whether the failure of GSC maintenance is due to differentiation without self-renewing divisions or due to aberrant divisions followed by degeneration, the ovarian development of the *piwi*<sup>1</sup> mutant was examined more closely. The *piwi*<sup>1</sup> third instar larval ovaries contain a normal number of 50-70 presumed GSCs that are mis-positioned. However, these GSCs still seem to, and indeed must have, partitioned during subsequent pupal development, since the number of ovarioles formed in most adult ovaries is within the normal range ( $13 \pm 3$  in mutant vs.  $15 \pm 3$  in wildtype). Hence, the effect of *piwi*<sup>1</sup> on GSC mis-positioning is not essential for oogenesis.

The main oogenic defect of the *piwi*<sup>1</sup> mutant is the differentiation of GSCs without self-renewing divisions immediately following the initiation of oogenesis. At this stage in wildtype ovaries, GSCs in 16-23 newly formed germaria have initiated asymmetric divisions to generate multiple developing germline cysts. However, in mutant ovaries, GSCs and their immediate differentiated daughters, cystoblasts, are undetectable, as indicated by the absence of spectrosome-containing germ cells. Instead, most ovaries contain differentiated germline cysts whose number approximately equals that of GSCs. These cysts are much larger in size and contain 2- to multi-cell stage fusomes, indicating their differentiating state. By the adult stage, most ovarioles contain only two normal or abnormal egg chambers derived from

these cysts, but no other germline cells. This defect contrasts the development of wildtype ovaries, in which ovarioles contain a fully developed germarium and a stage 1 egg chamber by the 48-hour pupal stage, and have produced multiple egg chambers by adult stage. This observation indicates that the mutant GSCs have differentiated into germline cysts without self-renewing divisions. The abnormal fusome morphology often seen in cysts suggest that they contain more or less than the normal number of 16 cells, which indeed was seen in adult egg chambers.

Thus, *piwi* is required for the self-renewing division of GSCs during oogenesis in *Drosophila*. Confocal images showing third instar larval ovaries from wild-type, *piwi*<sup>2</sup>, and *piwi*<sup>1</sup> mutant flies, respectively, were prepared and stained with a germline specific marker VASA in green and spectrin in red to mark somatic cells and spectrosomes/fusomes. GSCs were positioned medially along the anterioposterior axis in the wild-type and *piwi*<sup>2</sup> ovary; however, GSCs are mispositioned in the *piwi*<sup>1</sup> mutant ovary. Spectrosomes (Sp) appear as red dots in germ cells. A 50µm measurement bar was used in the analysis of the confocal images.

Confocal images of 24-hour pupal ovaries from wild-type and *piwi*<sup>1</sup> mutant flies, respectively, were prepared, with wild type stained solely for VASA while *piwi*<sup>1</sup> mutant was stained for both VASA in green and spectrin in red. At this stage, the wild-type ovary has partitioned into individual germaria (Ge), which contain GSCs and early germline cysts. In *piwi*<sup>1</sup> mutant ovaries germ cells still appear to have been partitioned, although the partition is distorted by the highly differentiated large germline cysts containing elaborate fusomes (Fu). The number of cysts did not exceed the number of GSCs.

Confocal images of 48-hour pupal ovaries from wild-type, *piwi*<sup>2</sup> and *piwi*<sup>1</sup> mutant flies, respectively, were prepared and stained with VASA in green and spectrin in red. By this stage, in both wild-type and *piwi*<sup>2</sup> mutant ovaries, germaria have fully developed and stage 1 egg chambers (S1) have budded off the germarium in synchrony due to continued GSC division. However, in *piwi*<sup>1</sup> mutant ovaries, the number of germline cysts remained unchanged, despite their continued differentiation.

A confocal image of a 72-hour pupal ovary from *piwi*<sup>2</sup> mutant flies was prepared and stained with VASA in green and spectrin stained in red. At this stage, GSCs fail to self-renew and some existing germline cysts start to degenerate. DAPI images of 0-1 day old adult ovarioles from wild-type, *piwi*<sup>2</sup> and *piwi*<sup>1</sup> mutant flies, respectively, were also prepared. Both *piwi*<sup>1</sup> (L) and *piwi*<sup>2</sup> (H) mutant ovarioles typically contain two normal or abnormal egg chambers connected to a germlineless germaria (Ge). In contrast, wild-type ovarioles contain a long string of developing egg chambers produced continuously by the germarium. *piwi*<sup>3</sup> phenotype is indistinguishable from that of *piwi*<sup>2</sup>.

Molecular cloning of *Drosophila piwi*. To study the molecular nature of the *piwi*-mediated mechanism, a 15 kb genomic region flanking the *piwi* insertions was cloned. Candidate transcripts in the region were identified, and a 6.8 kb genomic fragment (pRc12) containing a 3.1 kb transcript interrupted by *piwi* mutations (transcript c12) was able to rescue the *piwi* phenotype completely (Fig. 1). An overlapping 3.9 kb genomic fragment (pRLD) containing a 1.43 kb transcript immediately adjacent to the *piwi* mutations fails to rescue the *piwi* phenotype (Fig. 1). Therefore, the c12 transcript is the *piwi* mRNA.

*Drosophila* PIWI represents a novel class of conserved proteins. To study the structure of the *piwi* gene, two overlapping cDNA clones which together represent the 3.1 kb full length c12 mRNA were isolated and sequenced. The c12 cDNA sequence contains a large open reading frame (ORF) starting at +84 bp and encoding a protein of 843 amino acid residues with a predicted molecular weight of 97.2 kD (SEQ ID NOs:1 and 2). Multiple stop codons in all three frames are present upstream of the putative start codon. Downstream of the ORF is a 432 bp 3' untranslated region (UTR) followed by a 55-bp poly(A) tail not encoded in the genomic sequence. Thus, *piwi* RNA is present in *piwi*<sup>1</sup>/CyO adult males and females. Northern blots including female and male lanes containing approximately 10 and 1.5 mg of polyA<sup>+</sup> RNA, respectively, were prepared. They were exposed for 3.5 and 12 hours, respectively, and a band was observed at 3.1 kb in each lane.

The PIWI protein is a highly basic (pI 9.6) novel protein with no obvious similarity to other known proteins or functional motifs in the databases. It is characterized by alternating basic and acidic regions and is particularly basic over the C-terminal 100 amino acid residues. Hydropathy analysis indicates that the PIWI protein contains no significant local hydrophobic regions that could be potential signal peptide or transmembrane domains. PSORT analysis (Nakai and Kanehisa, 1992) predicts nuclear localization of the PIWI protein. The protein has 21 conserved protein kinase C phosphorylation sites, 14 casein kinase 2 phosphorylation sites, and 4 tyrosine phosphorylation sites, indicating its potential as a phosphorylation target.

To determine whether the PIWI protein is conserved during evolution, applicant searched for its homologous sequences at the protein level and identified two ORFs of unknown function from *C. elegans* and an expressed sequence tag (EST, AA430311) from a human testis cDNA library. Applicant isolated and sequenced cDNAs for the two *C. elegans* genes, herein named *prg-1* and *prg-2* (*prg* for *piwi*-related gene) to verify their homology to *Drosophila piwi*. The *prg-1* and *prg-2* genes share 40.1% and 38.5% amino acid identity to *piwi*, respectively, over their entire length. In the C-terminal 104 amino acid region, the homology increases to 55.8% and 56.7%, respectively. Moreover, *prg-1* and *prg-2* are 90% identical to each other over their full length and 98% identical at the C-terminus. This high degree of homology suggests that *prg-1* and *prg-2* may represent a gene duplication event. The two clones differ primarily in that *prg-1* is 60 amino acids longer at the N-terminus than *prg-2*. Using ACeDB (Thierry-Mieg and Durbin, 1992), *prg-1* was mapped to chromosome I between *unc-15* and *gld-1* in cosmid D2030 and *prg-2* was mapped to chromosome IV, between *unc-44* and *smg-7* on cosmid CO1G5.

To isolate human *piwi* homologs, the human EST (0.9 kb) clone was sequenced and used to screen a human testis cDNA library. A resulting 2.3 kb partial cDNA, herein named *hiwi* (for human *piwi*), shows 47.1% identical amino acid sequence to the *Drosophila piwi* over its full length, with 58.7% identity at the C-terminus. No *piwi*-related sequences were found from bacteria or yeast genomes whose entire sequences are known. This is consistent with

-87-

the stem cell-related function of *piwi* and indicative of *piwi* family-like genes specific for multicellular organisms.

In addition to the above *piwi* family homologs, a large number of putative and known proteins were identified from various animals and plants which share significant homology with *piwi* solely at their respective C-termini. This indicates that the *piwi* family is a large novel gene family. Within this family are 13 additional putative *C. elegans* proteins and three *A. thaliana* proteins, two of which are required for meristem cell divisions (Moussian, et al., 1998; Bohmert, et al., 1998). Particularly, between *piwi* and the three *Arabidopsis* genes, *ZWILLE* (*ZLL*) *argonaute* (*ago*), and *argonaute-like* (Moussian et al., 1998; Bohmert et al., 1998), an overall homology of approximately 20% amino acid identity was observed. The homology increases to 32% to 52% identity in four regions, 30-100 amino acid residues each, located throughout the length of the PIWI protein, including the highly conserved C-terminal region. Given the role of *ZLL* and *ago* in meristem cell division (Moussian et al., 1998; Bohmert et al., 1998), this homology may have important implications for a conserved stem cell mechanism.

PIWI, PRG-1, PRG-2, and HIWI differ from *ZLL* and *AGO* proteins, and especially from the 13 additional putative *C. elegans* proteins, predominantly at the N-terminus, suggesting that this region may be involved in *piwi*-specific function. The C-terminal conservation suggests that this region of PIWI may contain a novel functional domain that plays an important role for the general activity of these proteins in diverse biochemical processes, with the N-terminus rendering the specificity of the activity.

To examine the C-terminal region of homology more closely, the sequences were aligned using Block Maker™, which reveals characteristic regions of protein families (Henikoff et al., 1995) (Figure 2). Block Maker™ analysis identified a 43 amino acid domain conserved among all 22 proteins, within which five residues are absolutely conserved with defined spacing. Eight more residues are also conserved with defined spacing among all known genes across the phyla except for several *C. elegans* ORFs with unknown function. This region is referred to herein as the PIWI box and represents a

novel conserved functional motif. PIWI, that is the *piwi* family of gene products, thus represents a novel class of evolutionarily conserved proteins with conserved functionality, as described herein below.

*Drosophila piwi* mRNA is present in both the germline and soma during oogenesis. To investigate in which cells *piwi* is expressed to regulate GSC division, the expression pattern of *piwi* during oogenesis was examined by *in situ* hybridization of whole mount *Drosophila* ovaries with DIG-labeled DNA probes prepared from the *piwi* cDNA clone. *piwi* mRNA is detected specifically in the somatic terminal filament cells apical to GSCs in the germarium and anterior sheath cells as well as in the germline. In the germline, it is first abundantly expressed in region 2 of the germarium where 16-cell germline cysts are formed, persists at a lower level through stages 1-6 of oogenesis, is at its lowest level between stage 7-9, becomes strongly expressed again at stage 10, and eventually accumulates in early embryos (for staging, see King, 1970). Given that removing sheath cells does not affect oogenesis whereas the terminal filament cells play a role in regulating GSC division (Lin and Spradling, 1993), it is likely that the somatic expression in the terminal filament is responsible for *piwi* function in regulating GSC division.

Thus, *piwi* mRNA expression in *Drosophila* ovaries was characterized. A complete *Ore-R* ovariole was studied and showed germline *piwi* expression in region II of the germarium (Gell) and in early stage egg chambers up to stage 10 (S10) was studied. In the germarium, the *piwi* RNA was also detected in the terminal filament cells (TF) and epithelial sheath cells (ES) apical to GSCs. *piwi* RNA was also observed to be uniformly present in early embryos. The aforementioned structures were reviewed on slides, and the slides included bars denoting 50  $\mu$ m, 10  $\mu$ m and 50  $\mu$ m for ovariole, germarium and early embryos, respectively, to facilitate review.

Piwi expression in apical somatic cells regulates GSC division while its expression in the germline provides a maternal component for embryogenesis. To examine the roles of somatic and germline *piwi* expression, genetic clonal analyses on *piwi*<sup>1</sup> and *piwi*<sup>2</sup> mutations was conducted. First, the germline requirement of *piwi* was tested by germline clonal analysis using the FLP-DFS

technique (Chou and Perrimon, 1996). The *ovo<sup>D1</sup>* dominant mutation blocks oogenesis at the very beginning stage in a cell-autonomous manner, so that *piwi<sup>+</sup> ovo<sup>D1</sup>* females contain germ cells that never differentiate beyond the germarium. In this *ovo<sup>D1</sup>* background, applicant generated germline cells homozygous for either the *piwi<sup>1</sup>* or *piwi<sup>2</sup>* mutation by applying a heatshock treatment at the second and early third instar larval stage. The resulting adult females developed some completely normal ovarioles in which germline cells at all stages of oogenesis were observed, and mature eggs were continuously being produced (see Table 3).

This result demonstrates that removing the *piwi<sup>+</sup>ovo<sup>D1</sup>* chromosome from the germline allows oogenesis, including GSC division, to occur normally. Thus, the requirement of *piwi* for GSC division does not reside in the germline but in somatic cells.

Eggs produced from homozygous *piwi<sup>1</sup>* or *piwi<sup>2</sup>* GSC clones are arrested in embryogenesis, not rescuable by the paternal *piwi<sup>+</sup>* gene (see Table 3). Approximately 30% of the arrested embryos show severe mitotic defects during cleavage stage. The remaining embryos show various morphological defects during late embryogenesis, including a high frequency of severe deformation of the head region. These defects demonstrate that *piwi* expression in the germline provides an essential maternal contribution for embryogenesis.

To determine whether *piwi* is required in somatic cells outside the ovary for GSC maintenance, wildtype germaria were transplanted into the abdominal cavity of homozygous *piwi<sup>2</sup>* and wildtype females, which produced  $17 \pm 0$  egg chambers ( $n=13$ ) after 7 days of incubation in the host. Parallel transplantation into wildtype host yielded  $16 \pm 2$  egg chambers ( $n=5$ ), indicating that the wildtype germaria continued oogenesis at a normal rate in the *piwi<sup>2</sup>* females. Hence, *piwi* is not required in extra-ovarian cells for GSC maintenance.

To determine in which group of ovarian somatic cells *piwi* is required for GSC maintenance, applicant used a similar FRT-mediated technique for somatic clonal analysis of *piwi<sup>1</sup>* and *piwi<sup>2</sup>*, marking the *piwi<sup>+</sup>* chromosome with a transgenic myc tag so that homozygous *piwi<sup>-</sup>* somatic cells are identified by

their lack of myc epitope expression (Xu and Rubin, 1993). By inducing mitotic recombination at the second and third instar larval stages, *piwi*<sup>-</sup> somatic clones were generated throughout ovarioles, with some egg chambers completely covered by *piwi* mutant follicle cells. These egg chambers develop normally, indicating that *piwi* function is not required in follicle cells for egg chamber development. Since follicle cells are derived from their precursor cells in region II of the germarium (Margolis and Spradling, 1995), this suggests that *piwi* is not required in somatic cells from germarial region II on for GSC division and ovarian development. These results, together with the *piwi* expression pattern *in situ*, suggest that *piwi* is required in somatic cells in the anterior-most tip of the germarium to regulate GSC division.

Thus, *piwi* function is dispensable in the germline and differentiated follicle cells for GSC division. A pair of ovaries containing ovarioles with homozygous *piwi*<sup>-</sup> germline clones and ovarioles with *piwi*<sup>+</sup> *ovo*<sup>D1</sup> germline stained for DAPI to mark DNA and for VASA to mark germ cells for observation via confocal microscopy. The *piwi*<sup>+</sup> *ovo*<sup>D1</sup> germline cells were arrested at the beginning of oogenesis in the germarium due to the *ovo*<sup>D1</sup> mutation. By contrast, in ovarioles in the left ovary containing *piwi*<sup>-</sup> germline clones, GSCs divided normally, giving rise to a progression of morphologically wild-type egg chambers which eventually develop into mature eggs. This indicates that *piwi* is not required in the germline for GSC division and subsequent steps of oogenesis. To facilitate confocal microscopy, bars denoting 100 µm were employed in each slide.

Confocal microscopy was also employed to show that *piwi*<sup>-</sup> clones in differentiated follicle cells do not effect normal GSC division and egg chamber development. Egg chambers were stained with DAPI and a monoclonal anti-MYC antibody. Absence of MYC staining was indicative of homozygous *piwi*<sup>-</sup> clones. Induction of somatic clones around an entire early stage egg chamber had no effect on GSC division and oogenesis. Similarly, removing *piwi* function from some follicle cells of later stage egg chambers also had no effect on oogenesis. Therefore, *piwi* function is dispensable in differentiated follicle cells



for oogenesis including GSC division. To facilitate confocal microscopy, bars denoting 50  $\mu$ m were employed in each slide.

*piwi*-like genes in *C. elegans* are also required for germline self-renewal.

5 The high degree of sequence homology between *piwi* and its homologs in other organisms suggests a potential functional conservation. This hypothesis was tested in *C. elegans*. Guo and Kemphues (1995) have shown that the injection of specific antisense RNA into the germline syncytium of *C. elegans* eliminates maternal and zygotic gene activity, producing a gene-specific loss-of-function effect that may persist through several generations. This technique, as refined  
10 by Fire et al. (1998) and termed RNA-mediated interference (RNAi), was used to assess the function of *prg-1* and *prg-2*. Given the extremely high homology between *prg-1* and *prg-2*, an anti-*prg-1* RNA was used for injection to interfere with the function of both genes. The F1 progeny of the injected worms were designated as *prg-RNAi* worms for simplicity.

15 In wild-type *C. elegans*, two germline precursor cells, Z2 and Z3, give rise to approximately 2000 germ cells in the adult hermaphrodite. Germline proliferation occurs throughout most of larval development (L1-L4) and continues in the adult. This proliferation and maintenance of the germline requires signals from the DTC at the tip of each gonadal arm. During gonadal  
20 development, DTC migration results in the formation of two U-shaped gonadal arms by the L4 stage. Germline proliferation is limited to the distal end of each arm, forming a mitotic proliferation zone (MPZ) which serves as the GSC equivalent in *C. elegans*. Moving proximally, near the U turn of the gonad, germ cells enter meiotic prophase and then further differentiate into gametes  
25 at the proximal half of each arm, producing sperm in L4 and then oocytes in young adults.

The phenotype of *prg-RNAi* worms was examined both by quantitating their fertility and by assessing their gonadal and germline development via DAPI staining and Nomarski optics. As controls for RNA injection, worms were  
30 injected with BlueScript plasmid RNA sequences.

The *prg-RNAi* worms show three major aspects of germline defects. First, as indicated by the distribution of the number of progeny per F1 animal

(Fig. 3), the fertility difference between the *prg-RNAi* and control worms is highly significant ( $t$  value=7.93,  $p<0.0001$ ). On average, *prg-RNAi* worms produced  $92\pm 53$  progeny ( $n=99$ ), with 75.5% of *prg-RNAi* worms giving rise to fewer than 125 progeny. By contrast, control RNAi animals on average produced  $191\pm 40$  progeny ( $n=20$ ), with 90% of these animals giving rise to greater than 125 progeny. The significant reduction of the fertility, as well as other defects described below, may still only reflect a partial loss of *prg-1* and *prg-2* function, since the RNAi technique is known to phenocopy partial rather than complete loss-of-function mutants (Kuwabara, 1996).

Second, DAPI staining reveals a dramatic shortening of MPZ in *prg-RNAi* worms. On average, adult *prg-RNAi* worms exhibited a 50% reduction in the number of mitotic nuclei as compared to control animals. Associated with the reduction of the mitotic zone is a gonadal shortening. 57% of *prg-RNAi* worms ( $n=287$ ) exhibited a moderate to severe shortening. In the most severe case, the gonadal arm never made the U-turn. The number of sperm produced in these worms is greatly reduced as well. By contrast, only 7% of the BlueScript-RNA-injected control animals ( $n=548$ ) exhibited a mild gonadal shortening. The above observed gonadal defects in *prg-RNAi* worms indicate that *prg-1* and *prg-2* are essential for germline proliferation and maintenance. Since the gonadal shortening may be due to a defect in DTC migration, it is also possible that *prg-1* and *prg-2* play a role in proper gonadogenesis.

Thus, reducing *prg* activity results in reductions in both the MPZ and the mitotic index. Variable defects in both the size of the MPZ and the mitotic index were observed by DAPI staining and preparation of photomicrographs of same. The MPZ in each micrograph was denoted in white brackets. In moderate to severely defective animals, a 50% reduction in the number of mitotic nuclei and a concomitant five-fold reduction in the mitotic index were typically observed. A measurement bar was included in the wildtype photomicrograph and denoted 100  $\mu$ m for all panels.

In addition to the MPZ shortening, the mitotic index in the remaining mitotic zone is further reduced by 5.5 fold from  $(7.1\pm 2.2)\%$  ( $n=1,339$  nuclei,  $n_s=10$  worms) in the control RNAi injected worms to  $(1.3\pm 0.8)\%$  ( $n=1,409$

nuclei,  $n_s=10$  worms) in *prg-RNAi* worms with mildly to moderately shortened gonads. This indicates the important role of *prg-1* and *prg-2* in maintaining the mitotic ability of the germline stem cell nuclei.

*Drosophila piwi* is required for the self-renewing divisions of stem cells  
5 in the germline. The above experimentation demonstrates that *piwi* provides for female and male GSC maintenance predominantly by promoting the self-renewing division of GSCs. In a germarium, it normally takes at least 10 hours for a GSC division to occur and 5 days for a cystoblast to develop into a stage one egg chamber (King, 1970; Lin and Spradling, 1993). Necrosis and  
10 apoptosis during oogenesis are also easily detectable. Thus, examination of the *piwi* mutant phenotype at 24-hour intervals immediately prior to the onset of oogenesis is sufficient to monitor every major event in germline proliferation, differentiation, or cell death. Under these conditions, GSC divisions would have been detected as an increased number of germline cysts, whereas cell  
15 death process would have been easily detected by looking for pycnotic nuclei and necrotic bodies via DAPI staining and Nomarski optics (Lin and Spradling, 1993; Forbes et al., 1996a, b; Lin and Spradling, 1997). Numerous cases of GSC divisions and cell death in the germline during *Drosophila* oogenesis have been detected using these criteria (Lin and Spradling, 1993; Forbes et al.,  
20 1996a, b; Lin and Spradling, 1997).

Despite their high sensitivity, the techniques followed in this Example, even though more systematic, did not detect either GSC division or cell death in *piwi* mutants. Instead, immediately following the onset of oogenesis, only differentiating cysts were detected in *piwi*<sup>1</sup> mutant ovaries, with the cyst  
25 number generally corresponding to the number of GSCs. Although this does not rule out the possibility that infrequent GSC divisions or cell death might have been missed, which might account for the smaller-than-expected number of gametes sometimes observed in adult *piwi* mutants, these factors do not weaken the conclusion that the predominant defect of *piwi* mutants is the  
30 differentiation of GSCs without accompanying self-renewing division.

A number of elegant studies have identified genes that regulate germline proliferation. In mammalian systems, mechanisms mediated by paracrine

-94-

factors such as the *Steel* factor (the ligand of the *c-Kit* receptor tyrosine kinase), leukemia inhibitory factor (LIF), and basic fibroblast growth factor (bFGF), control the proliferation and survival of primordial germ cells, i.e., the precursors of GSCs (reviewed in Lin, 1997). In *C. elegans*, the LAG-2/GLP-1 pathway is essential for the self-renewal of a stem-cell equivalent population of germline nuclei in the gonad (reviewed in Kimble and Simpson, 1997). In the *Drosophila* ovary and testis, the BAM-mediated mechanism is involved in the further development of the differentiated GSC daughter (McKearin and Ohlstein, 1995); the *dpp* pathway plays a similar role in testis (Matunis et al. 1997). More recently, this pathway has been shown to be essential for GSC division and maintenance in the ovary (Xie and Spradling, 1998). Among these discoveries, *piwi* is the first gene known to mediate a somatic signaling mechanism to regulate the division and maintenance of GSCs in *Drosophila*. It therefore provides an important entry point for studying the somatic control of stem cell division in the germline.

The *piwi* gene family and the conservation of a stem cell self-renewing mechanism. The identification of *prg-1* and *prg-2* in *C. elegans* and *hiwi* in human suggest that *piwi* represents a novel class of evolutionarily conserved genes (described herein as the "*piwi* family") with potentially important functions in GSC self-renewal. Among this class of genes, the significantly higher homology between *piwi* and *hiwi* as compared to that between *piwi* and *prg-1/prg-2* suggests that *hiwi* function is closer to *piwi*. Consistent with this, GSC division and gametogenesis in humans are much more similar to that in *Drosophila* than that in *C. elegans* whose gonads contain syncytial mitotic germline nuclei which divide symmetrically and are capable of self-renewing only as a population. The RNAi experiments described herein, which likely interfered with the activity of both *prg-1* and *prg-2* due to their extremely high homology, caused germline depletion similar to that in *piwi* mutants. This suggests that the *piwi*-mediated mechanism in germline self-renewal is conserved even in this evolutionarily distant organism without stereotypic GSCs.

The conservation of the *piwi*-mediated mechanism appears to extend to the plant kingdom as well. A level of homology between PIWI and ZLL and AGO was observed. Intriguingly, ZLL is essential for maintaining stem cells of the shoot meristem in an undifferentiated state during the transition from embryo-specific development to repetitive organ formation through the self-perpetuating shoot meristem divisions (Moussian et al., 1998). AGO also plays an important role in maintaining normal apical shoot meristem function (Bohmert et al., 1998). Thus, the homology between *piwi* and *ZLL* and *ago* further suggests the existence of a family of novel genes essential for stem cell division in diverse organisms.

The conserved *piwi* mechanism may mediate cell-cell interactions. In *Drosophila*, genetic mosaic and *piwi* expression analyses together suggest that *piwi* function is required in the apical non-mitotic somatic cells to control GSC division. Similarly, in *A. thaliana*, it is thought that ZLL is required to maintain the undifferentiated state of shoot meristem stem cells by relaying positional information, possibly by mediating cell-cell interactions within the center of the shoot meristem (Moussian et al., 1998).

Stem cells are characterized by two common properties that extend across diverse species: the capacity for self-renewal and the ability to give rise to numerous progeny that are fated for further differentiation (for review see Lin, 1997; Morrison et al., 1997). Although significant progress has been made in identifying genes important for stem cell function, no common molecular mechanism shared by diverse stem cell types in diverse organisms has been characterized with respect to those two basic stem cell properties. The *piwi* gene family represents the first class of such genes. The analysis of the *piwi* gene family, therefore, provides an important first step towards the elucidation of molecular mechanisms underlying stem cell divisions.

The analyses of this Example were extended to homologs in *C. elegans*, an isolated human homolog (*hiwi*), and identified two *piwi*-like genes in *Arabidopsis* which are known to maintain meristem cells. This work reveals a new class of genes that play an essential role in stem cell division in multicellular organisms ranging from invertebrates to humans and plants.

-96-

Table 2

*piwi*<sup>1</sup> rescue results by pRc12 and pRLD<sup>a,b</sup>

	<i>piwi</i> <sup>1</sup> / <i>piwi</i> <sup>1</sup>	<i>piwi</i> <sup>1</sup> /CyO	<i>pRc12; piwi</i> <sup>1</sup> / <i>piwi</i> <sup>1</sup>	<i>pRLD; piwi</i> <sup>1</sup> / <i>piwi</i> <sup>1</sup>	
5	females	0 (136)	100 (45)	100 (39)	0 (81)
	males	0 (119)	97.8 (47)	88.4 (43)	0 (64)

<sup>a</sup>The percentage of fertile flies for each genotype and sex are indicated. The number of individuals tested for fertility (n) is in parentheses.

<sup>b</sup>Data are pooled from 6 and 4 independent insertion lines on the X and 3rd chromosomes for pRLD and pRc12, respectively. Not included are two other pRc12 rescue lines which showed sex-specific restoration of fertility, presumably due to position effects stemming from the insertion site.

Table 3

*piwi* function in GSC division is soma-dependent, while its function in embryogenesis is germline dependent\*

15	FLP/FRT <i>ovo</i> <sup>D1</sup> -induced germline clones	Heatshock	N	Avg. # eggs
	% embryos	laid/female	hatched	
	(a) <i>piwi</i> <sup>1</sup> , FRT/ <i>ovo</i> <sup>D1</sup> , FRT x <i>Oregon</i> <sup>R</sup> males +	37	50.4	0
	(b) <i>piwi</i> <sup>2</sup> , FRT/ <i>ovo</i> <sup>D1</sup> , FRT x <i>Oregon</i> <sup>R</sup> males +	26	66.7	0
20	(c) <i>piwi</i> <sup>1</sup> , FRT/ <i>ovo</i> <sup>D1</sup> , FRT x <i>Oregon</i> <sup>R</sup> males -	44	0	na
	(d) <i>piwi</i> <sup>2</sup> , FRT/ <i>ovo</i> <sup>D1</sup> , FRT x <i>Oregon</i> <sup>R</sup> males -	26	0	na
	(e) <i>ovo</i> <sup>D1</sup> , FRT/CyO x <i>Oregon</i> <sup>R</sup> males +	20	0	na
	(f) <i>ovo</i> <sup>D1</sup> , FRT/CyO x <i>Oregon</i> <sup>R</sup> males -	10	0	na

25 \* In all tester and control crosses, a number (N) of newly eclosed females were allowed to lay eggs at 25°C for 2 days. In both *piwi*<sup>1</sup> and *piwi*<sup>2</sup> testers, germline develops beyond that of *ovo*<sup>D1</sup> mutant ovarioles only following heat shock, giving rise to numerous eggs. This indicates the generation of *piwi*<sup>-</sup> germline clones and the dispensability of *piwi* function in the germline for

30 normal GSC division. However, none of the eggs develop beyond

-97-

embryogenesis, indicating the germline *piwi* expression is required as a maternal component for embryogenesis. The *ovo<sup>D1</sup>*, FRT control females lay no eggs either with or without heatshock. na=not applicable.

## Example 2

### 5                    Molecular Cloning and Expression Analysis of *Hiwi*.                     A Human Homolog of the *Drosophila Piwi* Gene Essential for                     Germline Stem Cell Division and Embryogenesis

10                    *piwi* is a novel gene first identified in *Drosophila* and was demonstrated to be essential for germline stem cell development in Example 1. In this Example the cloning of a 3.47 kb human *piwi* homolog that includes the whole open reading frame through cDNA library screening and 5' RACE is described. Sequence analysis showed high identity of HIWI to *Drosophila* PIWI, especially in the C-terminal domain of the protein. Chromosome mapping showed located *hiwi* to Chromosome 12q 24. 33, approximately 130kb from the AFM 295 ye5  
15                    marker. Northern blot analysis revealed that *hiwi* is only expressed in the testis among human adult tissues. The sequence and expression homology between human and *Drosophila piwi* genes suggested that *piwi* is conserved during evolution and plays an important role during early stages of spermatogenesis in human.

20                    Human spermatogenesis followed a precise and coordinated manner in the seminiferous epithelium of testis (reviewed in Belive, 1979; Thomas et al., 1989; Hoog, 1995). During this period the germ cells undergo a series of stage-specific morphological transitions, beginning with the proliferation and renewal of spermatogonia, followed by meiosis of spermatocytes, and finished by  
25                    differentiation of spermatids during spermiogenesis (Oakberg, 1956a; Nebel et al., 1961; Belive et al., 1977; Russel et al., 1990). Closing to the basal membrane, type A spermatogonia are stem cells (Bartmanska and Clermont, 1983; De Rooij, 1988), undergoing a series of mitoses to give rise to more type A cells or produce type B spermatogonia which passes through a mitotic  
30                    division to differentiate into spermatocytes.

Spermatocytes need to undergo a final round of DNA replication before entering meiosis. During meiosis, the synaptonemal complex and the recombination nodule become associated with chromosomes, which was thought to be important for two basic tasks in meiosis, the reduction of ploidy level and generation of new combinations of genes (reviewed in von Wettstein et al., 1984). The most radical morphological changes occurred during spermiogenesis (Oakberg, 1956b). Based on acrosome formation and nuclear condensation, the round spermid differentiated into mature spermatozoa which released into the lumen indicating the end of the spermatogenesis (reviewed in Bellve and O'Brien, 1983).

The proliferation and differentiation of germ cells in the seminiferous tubule has been demonstrated to be associated with adjacent somatic cells, the Sertoli cells (reviewed in Russel and Griswold, 1993; Griswold, 1995; Bitgood et al., 1996; de Miguel et al., 1997). The tubule of human testis is composed of germ cells in all stages of spermatogenesis. From basal membrane to lumen are cells ranging from stem cell to differentiated cell state and from proliferation to differentiation stage. Particular cell at particular developmental stage is associated with others at given sites in the tubule, the least mature stem cells are located furthest outside, while the most mature spermatozoa are located inside the seminiferous epithelium. The specific tubule organization and cellular association thus provide a good model system to study germ cell development.

Genes expressed in male germ cells had been identified in high mammals (Kramer and Erickson, 1982; Gold et al., 1983; Thomas et al., 1989; Wolgemuth and Watrin, 1991; Hoog, 1995). Some of these genes were expressed exclusively during spermatogenesis (reviewed in Willison and Ashworth, 1987; Erickson, 1990; Sandlow et al., 1997; van Rooijen et al., 1998), while others were expressed only in the later stages of spermatogenesis (Shankar et al., 1998; Zhu and Naz, 1998). However, very few genes had been demonstrated that was expressed specifically in the early stage of spermatogenesis, in the stem cells (Menke et al., 1997).



To study the mechanisms of *piwi* family function in germline stem cells in various organisms during evolution and especially its role in human beings, molecular cloning and expression analysis of human *piwi* family homolog was performed.

5        5'-RACE. A 2.3kb partial human *piwi* family cDNA clone was first isolated from adult human testis cDNA library through library screening. To obtain the 5' end of this gene, a cDNA synthesis primer and a gene specific primer has been designed. Adult human testis Poly A+ RNA was purchased from Clontech, Palo Alto, California. First-strand cDNA was obtained by using  
10       the above cDNA synthesis primer and AMV reverse transcriptase (Clontech, Palo Alto, California). Second-strand synthesis cDNA synthesis was performed with second-strand enzymes bought from Clontech, Palo Alto, California subsequently. Using the above gene specific primer and an adaptor primer obtained from Clontech, Palo Alto, California, an 1.4kb fragment has been  
15       obtained through 5'-RACE PCR by using Marathon™ cDNA Amplification Kit (Clontech, Palo Alto, California) and has been cloned into a pGEM-T Easy vector (Promega, Madison, Wisconsin) subsequently.

Construction of a full length human *piwi* homolog and its sequence analysis. The fully-length *hiwi* cDNA was generated in a modified pGEM-T  
20       Easy vector (Promega, Madison, Wisconsin) by partial digestion of above two overlapping cDNA fragments with *SacI* and subsequent ligation with T<sub>4</sub> DNA ligase (Gibco). The full length human *piwi* family homolog was sequenced fully on both strands using an ABI 377 Prism™ DNA Sequencer (Perkin-Elmer, Applied Biosystems) and analyzed using the DNASTar™ sequence analysis  
25       package (DNASTAR, Madison, Wisconsin). Homologous comparison between *hiwi* and its homologs was carried out using the Megalign™ program from the DNASTar™ package.

Chromosome mapping. Radiation Hybrid mapping against Stanford G3 Human Hamster panel was used for chromosome mapping for *hiwi*.

30       Northern Blot Analysis. A 0.9kb EST fragment used for Northern blot analysis was purchased from ATCC (American Type Culture Collection) and was subcloned into a pBlueScript KS(+) vector (Stratagene) with *EcoRI* and

*HindIII*. The 3.47kb plasmid DNA obtained from 5'RACE was used to make sense and antisense RNA probes for *in situ* hybridization. The above plasmid DNA were linearized with appropriate restriction enzymes. Using these linearized template DNA, sense and antisense RNA probes for Northern blot analysis and *in situ* hybridization were transcribed from the T7 and SP6 promoters with appropriate T7 and SP6 RNA polymerase (Gibco), following the protocols suggested by manufacturers. Probes used for Northern blot analysis were labeled by <sup>32</sup>P-UTP (Amersham).

Premade Northern blot of high-quality poly A+ RNA from different human tissues (Human Multiple Tissue Northern Blot II) was purchased from Clontech, Palo Alto, California. Human  $\beta$ -actin cDNA probe (Clontech, Palo Alto, California) was used as a positive control. ExpressHyb™ Hybridization Solution (Clontech, Palo Alto, California) was used as hybridization buffer for Northern analysis. The hybridization for both sample blot and positive control was performed under 68°C for 2 hours followed by 2 washes in 2x SSC, 0.05% SDS for 50 minutes at room temperature, and another 2 washes in 0.1x SSC, 0.1% SDS for 40 minutes at 50°C. The sample blot was exposed at -80°C for 10 hours while the positive control was exposed for 2 hours at the same temperature.

Cloning of full-length human *piwi* cDNA clone. An isolated 2.3 kb partial human *piwi* cDNA clone from a human testis library by library screening is described in Example 1. To obtain the 5' end of this gene, by using a gene-specific primer, a 1.4 kb cDNA fragment containing overlapping sequence was obtained through 5'RACE and subcloned. To obtain a full length human *piwi* family clone, the above overlapping cDNA clones were partially digested. Fragments of the correct sizes were subsequently ligated and subcloned into a pGEM-T Easy vector (Figure 4) to generate the 3472 bp full-length human *piwi* family (*hiwi*) that includes an open reading frame (ORF) encoding a protein of 861 amino acids, with a predicted molecular weight of 94.71kDa (Fig. 7 and Fig. 8). The ORF starts at 167bp and stops at 2749bp. Upstream of the start codon is a 166bp 5'untranslated region (5'UTR). Downstream of the stop codon is a 723bp 3'untranslated region (3'UTR).